

# A MODULAR ANALOGUE OF MOROZOV'S THEOREM ON MAXIMAL SUBALGEBRAS OF SIMPLE LIE ALGEBRAS

ALEXANDER PREMET

**ABSTRACT.** Let  $G$  be a simple algebraic group over an algebraically closed field of characteristic  $p > 0$  and suppose that  $p$  is a very good prime for  $G$ . In this paper we prove that any maximal Lie subalgebra  $M$  of  $\mathfrak{g} = \text{Lie}(G)$  with  $\text{rad}(M) \neq 0$  has the form  $M = \text{Lie}(P)$  for some maximal parabolic subgroup  $P$  of  $G$ . This means that Morozov's theorem on maximal subalgebras is valid under mild assumptions on  $G$ . We show that such assumptions are necessary by providing a counterexample to Morozov's theorem for groups of type  $E_8$  over fields of characteristic 5. Our proof relies on the main results and methods of the classification theory of finite dimensional simple Lie algebras over fields of prime characteristic.

## 1. INTRODUCTION

Let  $\mathbb{k}$  be an algebraically closed field of characteristic  $p > 0$  and let  $G$  be a simple algebraic group over  $\mathbb{k}$ . Let  $\Pi$  be a basis of simple roots of the root system  $\Phi$  of  $G$ . Recall that  $p$  is said to be *very good* for  $G$  if  $\Phi$  is not of type  $A_{kp-1}$  for some  $k \in \mathbb{Z}_{>0}$  and all coefficients  $n_\alpha$  of the highest root  $\tilde{\alpha} = \sum_{\alpha \in \Pi} n_\alpha \alpha \in \Phi^+(\Pi)$  are less than  $p$ . It is well known that if  $p$  is very good for  $\mathfrak{g}$  then the Lie algebra  $\mathfrak{g} = \text{Lie}(G)$  is simple and the centraliser of any semisimple element of  $\mathfrak{g}$  is a Levi subalgebra of  $\mathfrak{g}$ . The goal of this paper is to show that an analogue of Morozov's theorem on non-semisimple maximal subalgebras holds for the Lie algebra  $\mathfrak{g}$ . Over complex numbers, Morozov's theorem is the starting point of Dynkin's classification of the maximal subalgebras of simple Lie algebras [Dyn52] and one hopes that the theorem stated below will play a similar role in the modular setting.

**Theorem 1.1.** *Let  $G$  be a simple algebraic  $\mathbb{k}$ -group, where  $p = \text{char}(\mathbb{k})$  is a very good prime for  $G$ , and let  $M$  be a maximal Lie subalgebra of  $\mathfrak{g} = \text{Lie}(G)$  with  $\text{rad}(M) \neq 0$ . Then  $M = \text{Lie}(P)$  for some maximal parabolic subgroup  $P$  of  $G$ .*

In good characteristic, it follows from the description of maximal closed subsystems of irreducible root systems that if  $P$  is a maximal parabolic subgroup of  $G$  then  $\text{Lie}(P)$  is a maximal subalgebra of  $G$ ; see Lemma 2.4 for detail. Thus Theorem 1.1 characterises the class of non-semisimple maximal subalgebras of  $\mathfrak{g}$ .

If  $G$  is one of the classical groups  $\text{SL}(V)$ ,  $\text{SO}(V)$  or  $\text{Sp}(V)$  and  $p$  is a very good prime for  $G$  then Theorem 1.1 can be proved very quickly by using Lemma 2.1 (which holds when  $p$  is very good for  $G$ ) and the reducibility of the action of  $M$  on the  $G$ -module  $V$ ; see [HS15a, § 7] for a detailed argument. Therefore, in what follows we assume that  $G$  is an exceptional algebraic group. No good substitute for  $V$  is available in this case and our proof of Theorem 1.1 relies on completely different methods coming from the classification theory of finite dimensional simple Lie algebras over fields of characteristic  $p > 3$ . We shall see later that if  $p$  is not very good for  $G$  then Theorem 1.1 breaks

---

2010 *Mathematics Subject Classification.* 17B45.

Supported by the Leverhulme Trust (Grant PPG-2013-293).

down very badly for many simple algebraic groups  $G$  (classical and exceptional). For  $\mathfrak{g} = \mathfrak{sl}(V)$  with  $p \mid \dim V$  this is due to the fact that any maximal subalgebra of  $\mathfrak{g}$  acting irreducibly on  $V$  is neither semisimple nor parabolic because it has to contain the scalar endomorphisms of  $V$ .

In characteristic zero, Theorem 1.1 is a classical result of Lie Theory and it has a long history. It was first proved by Morozov in his doctoral dissertation published by Kazan State University in 1943. This text to which Dynkin refers in his seminal article [Dyn52] is hard to find nowadays. Nevertheless, Panyushev and Vinberg managed to do this and reported that Morozov's original proof was quite long and relied on case-by-case considerations; see [PV10]. Morozov was probably aware of that because in the short note [Mor56] he has simplified his original arguments. His shorter proof is easily accessible and is essentially reproduced in [Bou75, Ch. VIII, § 10]. It relies on some results and observations which are no longer valid in prime characteristic.

It should be stressed at this point that a group analogue of Theorem 1.1 holds without any restriction on  $p$ . More precisely, it was proved by Weisfeiler [Weis66] and independently by Borel–Tits [BT71, Corollaire 3.3] that if  $G$  is a connected reductive algebraic group defined over an algebraically closed field of arbitrary characteristic and  $H$  is a maximal subgroup of  $G$  then either  $H$  is reductive or  $H$  is a maximal parabolic subgroup of  $G$  (there are a couple of minor glitches in Weisfeiler's argument and it only works over perfect fields). Although this result implies Morozov's theorem in the characteristic zero case, it cannot be applied for proving Theorem 1.1 in full generality. Indeed, a priori it is not clear that  $M = \text{Lie}(H)$  for some Zariski closed subgroup  $H$  of  $G$  as all modular substitutes for exponentiation become feeble when  $p$  is small.

In the last section of the paper we exhibit two bizarre examples where the equality  $M = \text{Lie}(H)$  breaks down in bad characteristic (it would be interesting to determine all such instances). More precisely, in the case where  $G$  is of type  $E_8$  and  $p = 5$  we construct a 74-dimensional maximal Lie subalgebra  $\mathfrak{m}$  of  $\text{Lie}(G)$  which resides in the middle of a short exact sequence

$$0 \rightarrow A \rightarrow \mathfrak{m} \rightarrow W(2; \underline{1}) \rightarrow 0$$

where  $A = \text{nil}(\mathfrak{m})$  is an abelian ideal of  $\mathfrak{m}$  isomorphic to  $(O(2; \underline{1})/\mathbb{k}1)^*$  as  $W(2; \underline{1})$ -modules. Here  $O(2; \underline{1}) = \mathbb{k}[X, Y]/(X^5, Y^5)$ , a truncated polynomial ring in two variables, and  $W(2; \underline{1}) = \text{Der}(O(2; \underline{1}))$ , a Witt–Jacobson Lie algebra. Our example for groups  $G$  of type  $G_2$  is more straightforward and has to do with the fact that in characteristic 2 the Lie algebra  $\text{Lie}(G)$  forgets its identity and becomes isomorphic to  $\mathfrak{psl}_4$ .

Theorem 1.1 is the starting point of the joint project with David Stewart which aims to extend Dynkin's classification of maximal subalgebras to the case of reductive Lie algebras over fields of characteristic  $p > 3$ . At the end of the paper we put forward a conjecture on non-semisimple maximal subalgebras of Lie algebras of type  $E_8$  over fields of characteristic 5.

**Acknowledgement.** I would like to thank Jim Humphreys for useful comments on an earlier version of this paper and Tom Purslow for using GAP to verify some crucial properties of the subalgebra  $\mathfrak{m}$ . I am very grateful to David Stewart who has read the whole proof, suggested several improvements, pointed out a gap in an earlier version of (3.20) and sent me a long list of typos.

## 2. NOTATION AND PRELIMINARY RESULTS

2.1. Throughout this paper  $G$  is an exceptional algebraic  $\mathbb{k}$ -group and  $p = \text{char}(\mathbb{k})$  is a good prime for  $G$ . It is well known that in this case the Lie algebra  $\mathfrak{g} = \text{Lie}(G)$  is simple and its Killing form  $\kappa$  is non-degenerate. Being the Lie algebra of an affine algebraic group,  $\mathfrak{g}$  carries a canonical  $[p]$ -th

power map  $\mathfrak{g} \ni x \mapsto x^{[p]} \in \mathfrak{g}$  equivariant under the adjoint action of  $G$ . An element  $h \in \mathfrak{g}$  is called *toral* if  $h^{[p]} = h$ . The adjoint endomorphism of any toral element of  $\mathfrak{g}$  is semisimple and its eigenvalues lie in  $\mathbb{F}_p$ . Therefore, any such element is contained in a maximal toral subalgebra of  $\mathfrak{g}$ . Given  $x \in \mathfrak{g}$  we write  $\mathfrak{g}_x$  for the centraliser of  $x$  in  $\mathfrak{g}$ . If  $V$  is a vector space over  $\mathbb{k}$  of dimension  $rp$ , where  $r$  is a positive integer, then we denote by  $\mathfrak{psl}(V)$ ,  $\mathfrak{psl}_{rp}$  and  $\mathfrak{pgl}_{rp}$  the (restricted) Lie algebras  $\mathfrak{sl}(V)/\mathbb{k}\text{Id}_V$ ,  $\mathfrak{sl}_{rp}(\mathbb{k})/\mathbb{k}1$ , and  $\mathfrak{gl}_{rp}(\mathbb{k})/\mathbb{k}1$ , respectively. It is well known (and easily seen) that  $\mathfrak{psl}_{rp} = [\mathfrak{pgl}_{rp}, \mathfrak{pgl}_{rp}]$  has codimension 1 in  $\mathfrak{pgl}_{rp}$  and  $\mathfrak{pgl}_{rp} = \text{Lie}(\text{PGL}_{rp})$ . If  $(p, r) \neq (2, 1)$  then  $\mathfrak{psl}_{rp} \cong \mathfrak{psl}(V)$  is a simple Lie algebra.

Let  $\mathcal{N}(\mathfrak{g})$  denote the nilpotent cone of  $\mathfrak{g} = \text{Lie}(G)$ , the set of all  $x \in \mathfrak{g}$  with  $x^{[p]^e} = 0$  for  $e \gg 0$ . The group  $G$  acts on  $\mathcal{N}(\mathfrak{g})$  with finitely many orbits which are labelled by their weighted Dynkin diagrams exactly as in the characteristic 0 case; see [Pre03], for example. The dimensions of nilpotent orbits can be found in [Car93, pp. 401–407]. Let  $\mathcal{N}_p(\mathfrak{g}) = \{x \in \mathfrak{g} \mid x^{[p]} = 0\}$ , the *restricted nullcone* of  $\mathfrak{g}$ . By [CLNP3], the variety  $\mathcal{N}_p(\mathfrak{g})$  coincides with the Zariski closure of a single nilpotent  $G$ -orbit; in particular, it is always irreducible.

2.2. Given a Lie subalgebra  $L$  of  $\mathfrak{g}$  we denote by  $\text{nil}(L)$  the maximal ideal of  $L$  consisting of nilpotent elements of  $\mathfrak{g}$ .

**Lemma 2.1.** *If  $M$  is a maximal Lie subalgebra of  $\mathfrak{g}$  with  $\text{rad}(M) \neq 0$  then  $\text{nil}(M) \neq 0$ .*

*Proof.* Let  $A$  be a maximal abelian ideal of  $M$ . Since  $M$  is maximal in  $\mathfrak{g}$ , it is a restricted subalgebra of  $\mathfrak{g}$ . Since  $\text{rad}(M) \neq 0$  is a restricted ideal of  $M$ , the ideal  $A \neq 0$  is closed under taking  $[p]$ -powers in  $\mathfrak{g}$ . If  $A \subseteq \mathcal{N}(\mathfrak{g})$  then  $0 \neq A \subseteq \text{nil}(M)$  and we are done. So suppose  $A \not\subseteq \mathcal{N}(\mathfrak{g})$ . The  $A$  must contain a nonzero semisimple element of  $\mathfrak{g}$ , say  $t$ . Note that  $t$  lies in the span of all  $t^{[p]^i}$  with  $i \geq 1$ . As

$$[t^{[p]^i}, M] = (\text{ad } t)^{p^i}(M) \subseteq (\text{ad } t)^2(M) \subseteq [t, A] = 0$$

for all  $i \geq 1$ , this yields  $t \in \mathfrak{z}(M)$ . Therefore,  $M = \mathfrak{g}_t$  by the maximality of  $M$ . Since  $\mathfrak{g}$  is simple and  $p$  is a good prime for  $G$ , the centraliser  $\mathfrak{g}_t$  is a *proper* Levi subalgebra of  $\mathfrak{g}$ . However, such a subalgebra cannot be maximal as it normalises the nilradical of a proper parabolic subalgebra of  $\mathfrak{g}$ . This contradiction shows that  $\text{nil}(M) \neq 0$  as stated.  $\square$

2.3. Recall that a connected reductive  $\mathbb{k}$ -group  $G$  is called *standard* if the derived subgroup of  $G$  is simply connected,  $p$  is a good prime for  $G$ , and the Lie algebra  $\mathfrak{g} = \text{Lie}(G)$  admits a non-degenerate  $G$ -invariant symmetric bilinear form. The following lemma is a straightforward generalisation of the main result of [LMT09].

**Lemma 2.2.** *Let  $G$  be a standard reductive  $\mathbb{k}$ -group and let  $L$  be a Lie subalgebra of  $\mathfrak{g}$  such that  $[L, L]$  consists of nilpotent elements of  $\mathfrak{g}$ . The  $L$  is contained in a Borel subalgebra of  $\mathfrak{g}$ .*

*Proof.* Replacing  $L$  by its  $p$ -closure in  $\mathfrak{g}$  if need be we may assume that  $L$  is a restricted subalgebra of  $\mathfrak{g}$ . We use induction on the dimension of  $G$ . Suppose the statement holds for all standard reductive  $\mathbb{k}$ -groups of dimension  $\leq n$  (this is obviously true when  $n = 1$ ). Now suppose  $\dim(G) = n + 1$ . By Engel's theorem, the Lie algebra  $[L, L]$  is nilpotent. If  $[L, L] \neq 0$  then  $\text{nil}(L) \neq 0$ . If  $[L, L] = 0$  then  $L = L_s \oplus L_n$  where  $L_s$  is a toral subalgebra of  $\mathfrak{g}$  and  $L_n = \text{nil}(L)$ . If  $L = L_s$  then  $L_s \subseteq \text{Lie}(T)$  for some maximal torus  $T$  of  $G$ ; see [Hum67, Theorem 13.3]. So we are done in this case.

Thus we may assume that  $\mathfrak{n} := \text{nil}(L) \neq 0$ . Then  $\mathfrak{z}(\mathfrak{n}) \neq 0$ . As  $[L, L] \subseteq \text{nil}(L)$  by our assumption on  $L$ , the adjoint action of  $L$  induces a representation of the abelian Lie algebra  $L/[L, L]$  on  $\mathfrak{z}(\mathfrak{n})$ .

So there exists a nonzero  $e \in \mathfrak{z}(\mathfrak{n})$  such that  $[L, e] \subseteq \mathbb{k}e$ . As  $e$  is a nilpotent element of  $\mathfrak{g}$  it admits a cocharacter  $\lambda: \mathbb{k}^\times \rightarrow G$  optimal in the sense of the Kempf–Rousseau theory. Let  $P$  be the parabolic subgroup associated with  $\lambda$  and  $\mathfrak{p} = \text{Lie}(P)$ . Then  $\mathfrak{p} = \bigoplus_{i \geq 0} \mathfrak{g}(\lambda, i)$ . By [Pre03, Theorem 2.3], one can choose  $\lambda$  in such a way that  $e \in \mathfrak{g}(\lambda, 2)$  and  $\mathfrak{g}_e \subseteq \mathfrak{p}$ . Furthermore,  $[\mathfrak{g}(\lambda, i), e] = \mathfrak{g}(\lambda, i + 2)$  for all  $i \geq 0$ . Taking  $i = 0$  we find  $h_0 \in \mathfrak{g}(\lambda, 0)$  with  $[h_0, e] = e$ . This implies that

$$L \subseteq \mathfrak{n}_{\mathfrak{g}}(\mathbb{k}e) = \mathbb{k}h_0 \oplus \mathfrak{g}_e \subseteq \mathfrak{p}.$$

By our induction assumption, the image of  $L$  in the standard Lie algebra  $\mathfrak{l} := \mathfrak{p}/\text{nil}(\mathfrak{p})$  is contained in a Borel subalgebra of  $\mathfrak{l}$ , say  $\mathfrak{b}$ . Since the inverse image of  $\mathfrak{b}$  under the canonical homomorphism  $\mathfrak{p} \twoheadrightarrow \mathfrak{l}$  is a Borel subalgebra of  $\mathfrak{g}$ , this accomplishes the induction step of our proof.  $\square$

2.4. We denote by  $\mathcal{O}_{\min}$  the minimal nonzero nilpotent orbit in  $\mathfrak{g}$ . It consists of all nonzero  $e \in \mathfrak{g}$  with the property that  $[e, [e, \mathfrak{g}]] = \mathbb{k}e$ .

**Corollary 2.3.** *Let  $M$  be a maximal Lie subalgebra of  $\mathfrak{g}$  and denote by  $N$  the nilradical of  $M$ . Suppose  $N \neq 0$  and let  $R$  be any Lie subalgebra of  $M$  whose derived ideal  $[R, R]$  consists of nilpotent elements of  $\mathfrak{g}$ . Then the centraliser  $\mathfrak{c}_{\mathfrak{g}}(N)$  is an ideal of  $M$  and there exists  $e \in \mathfrak{c}_{\mathfrak{g}}(N) \cap \mathcal{O}_{\min}$  such that  $[R, e] \subseteq \mathbb{k}e$ .*

*Proof.* Since  $N$  is nilpotent we have that  $0 \neq \mathfrak{z}(N) \subseteq \mathfrak{c}_{\mathfrak{g}}(N)$ . Therefore,  $\mathfrak{c}_{\mathfrak{g}}(N)$  is a nonzero Lie subalgebra of  $\mathfrak{g}$ . If  $x \in M$ ,  $c \in \mathfrak{c}_{\mathfrak{g}}(N)$  and  $n \in N$  then  $[[x, c], n] = [x, [c, n]] - [c, [x, n]] = 0$ . So  $[M, \mathfrak{c}_{\mathfrak{g}}(N)] \subseteq \mathfrak{c}_{\mathfrak{g}}(N)$  which implies that  $\widetilde{M} := M + \mathfrak{c}_{\mathfrak{g}}(N)$  is a Lie subalgebra of  $\mathfrak{g}$ . Since  $\mathfrak{g}$  is a simple Lie algebra and  $\widetilde{M}$  normalises  $N$ , it must be that  $\widetilde{M} \neq \mathfrak{g}$ . As a result,  $\widetilde{M} = M$  forcing  $\mathfrak{c}_{\mathfrak{g}}(N) \subseteq M$ .

Let  $\widetilde{R} = R + N$ . Then it is immediate from Jacobson’s formula for  $p$ -th powers that  $[\widetilde{R}, \widetilde{R}] \subseteq [R, R] + N$  consists of nilpotent elements of  $\mathfrak{g}$ . By Lemma 2.2, there exists a Borel subalgebra  $\mathfrak{b}$  of  $\mathfrak{g}$  containing  $\widetilde{R}$ . Let  $B = T \cdot R_u(B)$  be the Borel subgroup of  $G$  such that  $\mathfrak{b} = \text{Lie}(B)$ . The maximal torus  $T$  of  $B$  preserves the centre  $\mathfrak{z}(\mathfrak{n}_+)$  of  $\mathfrak{n}_+ := \text{Lie}(R_u(B))$ . This implies that  $\mathfrak{z}(\mathfrak{n}_+)$  is spanned by root vectors relative to  $T$ . Since  $p$  is a good prime for  $G$  and  $\mathfrak{n}_+$  contains all simple root vectors with respect to  $T$ , this yields that  $\mathfrak{z}(\mathfrak{n}_+) = \mathbb{k}e$  where  $e$  is a highest root vector of  $\mathfrak{n}_+$ . It is well known (and easily seen) that the latter belongs to  $\mathcal{O}_{\min}$ . As  $R \subseteq \widetilde{R} \subseteq \mathfrak{b}$  and  $\mathfrak{b}$  normalises  $\mathfrak{z}(\mathfrak{n}_+)$  it must be that  $[R, e] \subseteq \mathbb{k}e$ . Since  $N \subseteq \mathfrak{n}_+$  we also have that  $e \in \mathfrak{c}_{\mathfrak{g}}(N)$ . This completes the proof.  $\square$

2.5. A Lie subalgebra of  $\mathfrak{g}$  is called *regular* if it contains a maximal toral subalgebra of the restricted Lie algebra  $\mathfrak{g}$ .

**Lemma 2.4.** *Let  $M$  be a maximal subalgebra of  $\mathfrak{g}$  with  $\text{nil}(M) \neq 0$ . If  $M$  is regular then it is a parabolic subalgebra of  $\mathfrak{g}$ .*

*Proof.* Let  $\mathfrak{t}$  be a maximal toral subalgebra of  $\mathfrak{g}$  contained in  $M$ . It follows from [Bor91, 11.8] that there exists a maximal torus  $T$  in  $G$  such that  $\mathfrak{t} = \text{Lie}(T)$  (see also [Hum67, Theorem 13.3]). Since  $p > 3$ , the toral subalgebra  $\mathfrak{t}$  is a classical Cartan subalgebra of  $\mathfrak{g}$  in the sense of Seligman and the Lie algebra  $\mathfrak{g}$  satisfies the Seligman–Mills axioms; see [Sel67, Ch. II, § 3]. In particular, this means that all root spaces of  $\mathfrak{g}$  with respect to  $\mathfrak{t}$  are 1-dimensional. As a consequence,  $M$  is  $(\text{Ad } T)$ -stable.

Let  $\Phi$  be the root system of  $G$  with respect to  $T$ . Given  $\alpha \in \Phi$  we denote by  $\mathfrak{g}_{\alpha}$  the root subspace of  $\mathfrak{g}$  with respect to  $T$ . Then  $\mathfrak{g}_{\alpha} = \mathbb{k}e_{\alpha}$  for some root element  $e_{\alpha} \in \mathfrak{g}$ . The preceding remark entails that there exists a subset  $\Psi$  of  $\Phi$  such that  $M = \mathfrak{t} \oplus \sum_{\alpha \in \Psi} \mathbb{k}e_{\alpha}$ . Moreover, it follows

from [Sel67, Ch. II, § 4], for example, that if  $\alpha, \beta \in \Psi$  and  $\alpha + \beta \in \Phi$  then  $\alpha + \beta \in \Psi$ . In other words,  $\Psi$  is a closed subset of  $\Phi$  in the sense of [Bou68, Ch. VI, § 1, Sect. 7].

Let  $\mathfrak{g}_{\mathbb{Z}}$  be the Chevalley  $\mathbb{Z}$ -form of  $\mathfrak{g}$  associated with the root system  $\Phi$ . The above discussion shows that  $M$  is obtained by base change from the  $\mathbb{Z}$ -subalgebra  $\mathfrak{g}_{\mathbb{Z}}(\Psi)$  of  $\mathfrak{g}_{\mathbb{Z}}$  spanned by all root vectors  $E_{\alpha} \in \mathfrak{g}_{\mathbb{Z}}$  with  $\alpha \in \Psi$  and all commutators  $[E_{\beta}, E_{-\beta}]$  with  $\beta \in \Phi$ . The maximality of  $M$  in  $\mathfrak{g}$  now implies that  $\mathfrak{g}_{\mathbb{Z}}(\Psi) \otimes_{\mathbb{Z}} \mathbb{C}$  is a maximal subalgebra of  $\mathfrak{g}_{\mathbb{C}} := \mathfrak{g}_{\mathbb{Z}} \otimes_{\mathbb{Z}} \mathbb{C}$ . Then Morozov's theorem on maximal subalgebras of  $\mathfrak{g}_{\mathbb{C}}$  yields that the closed subset  $\Psi$  of  $\Phi$  is either symmetric or parabolic. If  $\Psi$  is symmetric then it is not hard to see that the restriction of the Killing form of  $\mathfrak{g}$  to  $M = \mathfrak{g}_{\mathbb{Z}}(\Psi) \otimes_{\mathbb{Z}} \mathbb{k}$  is non-degenerate. This, however, contradicts our assumption that  $\text{nil}(M) \neq 0$ . Therefore,  $\Psi$  must be a parabolic subset of  $\Phi$ . This completes the proof.  $\square$

2.6. Given a finite dimensional semisimple Lie algebra  $L$  we write  $L_p$  for the  $p$ -closure of  $L \cong \text{ad } L$  in  $\text{Der}(L)$ . This semisimple restricted Lie algebra is often referred to as the *minimal  $p$ -envelope* of  $L$ . To ease notation we shall often identify  $L$  with  $\text{ad } L \subseteq \text{Der}(L)$ . If  $L = \bigoplus_{i \in \mathbb{Z}} L_i$  is a graded Lie algebra then  $\text{Der}(L)$  has a natural grading, too, and Jacobson's formula for  $p$ -th powers implies that  $L_p$  is spanned by the  $p$ -closure of  $L_0$  in  $\text{Der}(L)$  and all  $L_i^{p^j}$  with  $i \neq 0$  and  $j \in \mathbb{Z}_{\geq 0}$ . Recall that an  $L$ -module  $V = \bigoplus_{i \in \mathbb{Z}} V_i$  is called *graded* if  $L_i \cdot V_j \subseteq V_{i+j}$  for all  $i, j \in \mathbb{Z}$ . We say that  $V$  is *restrictable* if there exists a restricted representation  $\rho: L_p \rightarrow \mathfrak{gl}(V)$  such that  $(\rho(x))(v) = x.v$  and for all  $x \in L$  and  $v \in V$ .

**Lemma 2.5.** *Let  $L$  be a finite dimensional semisimple graded Lie algebra over  $\mathbb{k}$  and suppose that its zero component  $L_0$  is a restricted subalgebra of  $L_p$ . Let  $V$  be a finite dimensional graded  $L$ -module, write  $\pi$  for the corresponding representation of  $L$ , and suppose that  $\pi|_{L_0}$  is a restricted representation of  $L_0$ . Then all composition factors of the  $L$ -module  $V$  are graded and restrictable.*

*Proof.* (a) We first show by induction on composition length  $l$  of  $V$  that all composition factors of  $V$  are graded. If  $V$  is irreducible then, of course, there is nothing to prove. So suppose  $V$  is reducible and let  $L_{\pm} = \bigoplus_{i > 0} L_{\pm i}$ . These graded Lie subalgebras of  $L$  are  $(\text{ad } L_0)$ -stable and act nilpotently on  $V$ . It follows that  $V_+ := \{v \in V \mid L_+ \cdot v = 0\}$  is a nonzero graded subspace of  $V$  and each graded component of  $V_+$  is invariant under the action of  $L_0$ . Let  $r$  be the smallest integer such that  $V_+ \cap V_r \neq 0$  and put  $W := U(L_-) \cdot (V_+ \cap V_r)$ . It is straightforward to see that  $W$  is a graded subspace of  $V$  invariant under the action of  $L_{\pm}$  and  $L_0$ . Our choice of  $r$  then implies that  $W$  is an irreducible  $L$ -submodule of  $V$ . As a consequence, the quotient module  $V/W$  has a natural structure of a graded  $L$ -module. Since the length of  $V/W$  is smaller than that of  $V$  and the composition factors of  $V$  are independent of the choice of a composition series (up to isomorphism), the statement follows by induction on  $l$ .

(b) Next we show that each composition factor of  $V$  is restrictable. In view of part (a) no generality will be lost by assuming that  $V$  is an irreducible  $L$ -module. Let  $\mathcal{L} = \sum_{i \geq 0} L^{p^i}$  be the  $p$ -envelope of  $L$  in the universal enveloping algebra  $U(L)$ . This is an infinite dimensional restricted Lie algebra with an enormous centre  $\mathfrak{z}(\mathcal{L})$ . The action of  $U(L)$  gives  $V$  a natural structure of a restricted  $\mathcal{L}$ -module. Let  $\rho: \mathcal{L} \rightarrow \mathfrak{gl}(V)$  denote the corresponding representation of  $\mathcal{L}$ . Since  $V$  is an irreducible  $L$ -module,  $\mathfrak{z}(\mathcal{L})$  acts on  $V$  by scalar linear operators. Now define  $\mathfrak{z}'(\mathcal{L}) := \mathfrak{z}(\mathcal{L}) \cap (\ker \rho)$  and put  $\tilde{\mathcal{L}} := \mathcal{L}/\mathfrak{z}'(\mathcal{L})$ . As  $L$  is semisimple,  $\tilde{\mathcal{L}}$  is a  $p$ -envelope of  $L$  and its centre  $\mathfrak{z} = \mathfrak{z}(\mathcal{L})/\mathfrak{z}'(\mathcal{L})$  has dimension  $\leq 1$ . Moreover,  $V$  is a restricted  $\tilde{\mathcal{L}}$ -module and  $\mathcal{L}/\mathfrak{z}(\mathcal{L}) \cong L_p$  as restricted Lie algebras; see [Str04, Theorem 1.1.7]. We denote by  $x \mapsto x^{[p]}$  the  $p$ -power map of  $\tilde{\mathcal{L}}$  induced by that of  $\mathcal{L}$ .



Abusing notation, we denote by  $\rho$  the representation of  $\tilde{L}_p$  in  $\mathfrak{gl}(V)$ . Then  $\rho(x^{[p]}) = \rho(x)^p$  for all  $x \in \tilde{L}_p$ .

Suppose  $\mathfrak{z} = \mathbb{k}z$  for some nonzero  $z \in \mathfrak{z}$ . Then we may assume further that  $\rho(z) = \text{Id}_V$ . As  $\tilde{L}_p$  is a  $p$ -envelope of  $L$ , we may identify  $L$  with a Lie subalgebra of  $\tilde{L}_p$  such that  $\tilde{L}_p = \sum_{i \geq 0} L^{[p]^i}$ . There exists a vector space complement  $L'_p$  to  $\mathfrak{z}$  in  $\tilde{L}_p$  which contains  $L$ . Then  $L'_p$  is a Lie subalgebra of  $\tilde{L}_p$  isomorphic to  $L_p$  and hence carries a  $p$ -th power map  $x \mapsto x^p$  induced by that of  $L_p$ . It is easy to see that  $x^{[p]} - x^p = \chi(x)^p z$  for some linear function  $\chi \in L^*$ . Since  $\rho|_L = \pi$  and  $V$  is a graded  $L$ -module, the subalgebras  $L_\pm$  act nilpotently on  $V$ . If  $x \in L_\pm$  then  $\rho(x^{[p]^N}) = \rho(x)^{p^N} = \pi(x)^{p^N} = 0$  for  $N \gg 0$  since  $\rho$  is a restricted representation of  $\tilde{L}_p$  and  $x^{p^N} = 0$  (recall that  $L_p \subseteq \text{Der}(L)$ ). Since  $[\rho(x^{[p]}), \rho(x^p)] = 0$ , the linear operator  $\rho(x^{[p]} - x^p) = \rho(x^{[p]}) - \rho(x^p)$  is nilpotent. As  $\rho(x^{[p]} - x^p) = \chi(x)^p \cdot \text{Id}_V$  we deduce that  $\chi$  vanishes on  $L_\pm$ . If  $x \in L_0$  then  $\pi(x^p) = \pi(x)^p = \rho(x)^p$  by our assumption on  $\pi$ . But then  $\rho(x^{[p]} - x^p) = \rho(x)^p - \rho(x)^p = 0$ , yielding  $\chi|_{L_0} = 0$ . As a result,  $\chi = 0$  and hence  $x^{[p]} = x^p$  for all  $x \in \tilde{L}$ . Since  $L'_p$  is the  $p$ -envelope of  $L$  under our second  $p$ -mapping  $x \mapsto x^p$ , we now obtain that  $\tilde{L}_p = L'_p$ . Since this contradicts our assumption that  $\mathfrak{z} \neq 0$  we now deduce that  $\tilde{L}_p$  is the minimal  $p$ -envelope of  $L$ . Then  $\tilde{L}_p \cong L_p$  as restricted Lie algebras, see [Str04, Corollary 1.1.8(2)].  $\square$

2.7. In order to deal with Hamiltonian Lie algebras which will come to complicate things in the next section we must look very closely at the  $G$ -orbits of certain special toral elements of  $\mathfrak{g}$ .

**Definition 2.6.** Let  $d$  be a positive integer. A toral element  $h \in \mathfrak{g}$  is said to be *d-balanced* if  $\dim \mathfrak{g}(h, i) = \dim \mathfrak{g}(h, j)$  for all  $i, j \in \mathbb{F}_p^\times$  and all eigenspaces  $\mathfrak{g}(h, i)$  with  $i \neq 0$  have dimension divisible by  $d$ .

If  $h$  is a  $d$ -balanced toral element of  $\mathfrak{g}$  then

$$(1) \quad \kappa(h, h) = \text{tr}(\text{ad } h)^2 = \sum_{i=1}^{p-1} \dim \mathfrak{g}(h, i) \cdot i^2 = \dim \mathfrak{g}(h, 1) \cdot (p-1)p(2p-1)/6 = 0.$$

**Proposition 2.7.** Let  $G$  be an exceptional algebraic  $\mathbb{k}$ -group, where  $p = \text{char}(\mathbb{k})$  is a good prime for  $G$ , and let  $h$  be a nonzero  $p$ -balanced toral element of  $\mathfrak{g} = \text{Lie}(G)$ . Then one of the following cases occurs:

- (i)  $G$  is of type  $E_6$ ,  $p = 5$ , the root system of  $G_h$  has type  $A_3$ , and  $\dim \mathfrak{g}_h = 18$ ;
- (ii)  $G$  is of type  $E_7$ ,  $p = 5$ , the root system of  $G_h$  has type  $D_4A_1$ , and  $\dim \mathfrak{g}_h = 33$ ;
- (iii)  $G$  is of type  $E_8$ ,  $p = 7$ , the root system of  $G_h$  has type  $E_6$ , and  $\dim \mathfrak{g}_h = 80$ ;
- (iv)  $G$  is of type  $E_8$ ,  $p = 7$ , the root system of  $G_h$  has type  $D_4A_2$ , and  $\dim \mathfrak{g}_h = 38$ ;
- (v)  $G$  is of type  $E_8$ ,  $p = 11$ , the root system of  $G_h$  has type  $A_4$ , and  $\dim \mathfrak{g}_h = 28$ .
- (vi)  $G$  is of type  $F_4$ ,  $p = 7$ , the root system of  $G_h$  has type  $A_2$ , and  $\dim \mathfrak{g}_h = 10$ .

Furthermore, in all cases except (vi) the nonzero  $p$ -balanced toral elements of  $\mathfrak{g}$  are conjugate under the adjoint action of  $G$ , whereas in case (vi) there are two  $p$ -balanced  $G$ -orbits.

*Proof.* (a) We may assume without loss of generality that  $G$  is a group of adjoint type. Let  $\Phi$  be the root system of  $G$  with respect to a maximal torus  $T$  of  $G$  and let  $\Pi = \{\alpha_1, \dots, \alpha_\ell\}$  be a basis of simple roots of  $\Phi$ , so that  $\ell = \dim T$ . Let  $\Phi^+$  be the positive system associated with  $\Pi$ . Any root  $\gamma \in \Phi^+$  can be uniquely expressed as  $\gamma = \sum_{i=1}^\ell \nu_i(\gamma) \alpha_i$  for some  $\nu_i(\gamma) \in \mathbb{Z}_{\geq 0}$ . In what follows we always use Bourbaki's numbering of simple roots in  $\Pi$ ; see [Bou68, Planches I–IX].

Since  $h$  is contained in a maximal toral subalgebra of  $\mathfrak{g}$  and all such subalgebras are  $(\text{Ad } G)$ -conjugate, we may assume that  $h \in \text{Lie}(T)$ ; see [Bor91, 11.8] of [Hum67, Theorem 13.3]. Then  $\mathfrak{g}_h$  is  $(\text{Ad } T)$ -stable. Since  $p$  is a good prime for  $G$  we may assume further that  $\mathfrak{g}_h$  is the standard Levi subalgebra of  $\mathfrak{g}$  associated with a subset  $\Pi_0$  of  $\Pi$ . Furthermore, there exist  $t_1, \dots, t_\ell \in \mathfrak{t}$  such that  $(d\alpha_i)(t_j) = \delta_{i,j}$  for all  $1 \leq i, j \leq \ell$ . These elements form a  $\mathbb{k}$ -basis of  $\mathfrak{t}$ . As a consequence, the centre of  $\mathfrak{g}_h$  has a  $\mathbb{k}$ -basis consisting of all  $t_i$  with  $1 \leq i \leq \ell$  such that  $(d\beta)(t_i) = 0$  for all  $\beta \in \Pi_0$ .

Let  $\Psi$  be the root system of  $\mathfrak{g}_h$  relative to  $T$  and put  $\Psi^+ := \Psi \cap \Phi^+$ . Then  $\dim \mathfrak{g}_h = \ell + 2|\Psi^+|$  and

$$(2) \quad \dim \mathfrak{g} - \dim \mathfrak{g}_h = \sum_{i \in \mathbb{F}_p^\times} \dim \mathfrak{g}(h, i) \equiv 0 \pmod{p(p-1)},$$

by our assumption on  $h$ . Hence  $|\Phi^+| \equiv |\Psi^+| \pmod{p(p-1)/2}$ .

(b) Let  $G_{\mathbb{C}}$  be the complex group of adjoint type with root system  $\Phi$ . We may assume without loss of generality that  $T$  is obtained by base change from a maximal torus  $T_{\mathbb{C}}$  in  $G_{\mathbb{C}}$ . In what follows we are going to rely on the well known analogy between toral elements of  $\mathfrak{g}$  and elements of order  $p$  in  $G_{\mathbb{C}}$ ; see [Ser06]. This will enable us to use extended Dynkin diagrams and Kac coordinates for labelling toral  $G$ -orbits in  $\mathfrak{g}$ .

More precisely, let  $\zeta$  be a primitive  $p$ -th root of unity in  $\mathbb{C}$ . Any toral element of  $\mathfrak{t}$  is a linear combination of semisimple root vectors  $(d\beta^\vee)(1) \in \text{Lie}(\beta^\vee(k^\times)) \subset T$  with coefficients in  $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ . If  $i \in \mathbb{F}_p^\times$  then  $\mathfrak{g}(h, i)$  is spanned the root spaces  $\mathfrak{g}_\gamma$  with  $\gamma(h) = i$ . Let  $\mathfrak{g}_{\mathbb{Z}} = \mathfrak{t}_{\mathbb{Z}} \oplus (\bigoplus_{\gamma \in \Phi} \mathbb{Z}e_\gamma)$  be an admissible  $\mathbb{Z}$ -lattice in  $\mathfrak{g}_{\mathbb{C}}$  such that  $\mathfrak{t} = \mathfrak{t}_{\mathbb{Z}} \otimes_{\mathbb{Z}} \mathbb{k}$  and  $\mathfrak{g}_\gamma = \mathbb{k}(e_\gamma \otimes 1)$  for all  $\gamma \in \Phi$ . Since  $G_{\mathbb{C}}$  is a group of adjoint type there exists a unique  $\sigma = \sigma(h) \in T_{\mathbb{C}}$  such that  $\sigma^p = 1$  and

$$\mathfrak{g}_{\mathbb{C}}(\sigma, \zeta^i) = \mathfrak{g}_{\mathbb{Z}}(\sigma, \zeta^i) \otimes_{\mathbb{Z}} \mathbb{C} \quad \text{and} \quad \mathfrak{g}(h, i) = \mathfrak{g}_{\mathbb{Z}}(\sigma, \zeta^i) \otimes_{\mathbb{Z}} \mathbb{k} \quad (\forall i \in (\mathbb{Z}/p\mathbb{Z})^\times)$$

where  $\mathfrak{g}_{\mathbb{C}}(\sigma, \zeta^i)$  is the  $\zeta^i$ -eigenspace of  $\sigma$  and  $\mathfrak{g}_{\mathbb{Z}}(\sigma, \zeta^i)$  is  $\mathbb{Z}$ -span of all  $e_\gamma$  with  $\gamma(\sigma) = \zeta^i$ .

Let  $P(\Phi^\vee)$  be the weight lattice of the dual root system  $\Phi^\vee$  and let  $W = N_G(T)/T \cong N_{G_{\mathbb{C}}}(T_{\mathbb{C}})/T_{\mathbb{C}}$  be the Weyl group of  $G$ . Since  $G_{\mathbb{C}}$  is a group of adjoint type the lattice of cocharacters  $T_{\mathbb{C}}$  coincides with  $P(\Phi^\vee)$  which implies that the  $\mathbb{F}_p W$ -module of all elements of order  $p$  in  $T_{\mathbb{C}}$  identifies with  $P(\Phi^\vee) \otimes_{\mathbb{Z}} \mathbb{F}_p$ . Since  $p > 3$  the latter space is  $W$ -equivariantly isomorphic to  $\mathfrak{t}^{\text{tor}}$ , the  $\mathbb{F}_p$ -subspace of all toral elements of  $\mathfrak{t}$ . From this it follows that there is a natural bijection between the toral  $G$ -orbits in  $\mathfrak{g}$  and the conjugacy classes of elements of order  $p$  in  $G_{\mathbb{C}}$ ; see [Ser06] for more detail.

Let  $\tilde{\alpha} = \sum_{\alpha \in \Pi} n_\alpha \alpha$  be the highest root in  $\Phi^+$  and put  $\tilde{\Pi} := \Pi \sqcup \{\alpha_0\}$  where  $\alpha_0 = -\tilde{\alpha}$ . To any nonzero collection of non-negative integers  $\mathbf{a} = (a_\alpha \mid \alpha \in \Pi)$  such that  $a_\alpha < p$  for all  $\alpha \in \Pi$  and  $\sum_{\alpha \in \Pi} a_\alpha n_\alpha \leq p$  we attach a unique element  $\sigma_{\mathbf{a}} \in T_{\mathbb{C}}$  of order  $p$  by imposing that  $\sigma_{\mathbf{a}}(e_\alpha) = \zeta^{a_\alpha} e_\alpha$  for all  $\alpha \in \Pi$  (here  $e_\gamma$  stands for a root vector of  $\mathfrak{g}_{\mathbb{C}}$  corresponding to  $\gamma \in \Phi$ ). It is well known that any element of order  $p$  in  $G$  is conjugate to one of the  $\sigma_{\mathbf{a}}$ 's. If  $\sigma$  is conjugate to  $\sigma_{\mathbf{a}}$  then the collection  $(a_\alpha \mid \alpha \in \tilde{\Pi}) = (a_{\alpha_0}, \mathbf{a})$  with  $a_{\alpha_0} = p - \sum_{\alpha \in \Pi} a_\alpha n_\alpha$  is sometimes referred to as *Kac coordinates* of  $\sigma$ . It is known that two different Kac coordinates represent the same conjugacy class in  $G_{\mathbb{C}}$  if and only if one can be obtained from the other by applying a suitable element from the group  $\text{Stab}_W(\tilde{\Pi}) \cong Z(\tilde{G}_{\mathbb{C}})$  (here  $\tilde{G}_{\mathbb{C}}$  stands for the simple simply connected algebraic group over  $\mathbb{C}$  with root system  $\Phi$ ). The action of  $\text{Stab}_W(\tilde{\Pi})$  on  $\tilde{\Pi}$  is described in [Bou68, Planches I–IX] (recall that our numbering of simple roots is compatible with that of *loc. cit.*). It is immediate from the above that the set  $\{\alpha \in \tilde{\Pi} \mid a_\alpha = 0\}$  forms a basis of simple roots of the fixed-point algebra  $\mathfrak{g}^{\sigma_{\mathbf{a}}} = \{x \in \mathfrak{g}_{\mathbb{C}} \mid \sigma_{\mathbf{a}}(x) = x\}$ . To ease notation we set  $a_i := a_{\alpha_i}$  for  $0 \leq i \leq \ell$ .

(c) In order to reduce the number of cases that we have to investigate closely we shall rely on the theory of sheets in  $\mathfrak{g}$  as presented in [PSt15]. Put  $\mathfrak{l} := \mathfrak{g}_h$  and let  $L$  be the standard Levi

subgroup of  $G$  with  $\mathfrak{l} = \text{Lie}(L)$ . Write  $L_{\mathbb{C}}$  for the standard Levi subgroup of  $G_{\mathbb{C}}$  associated with the subset  $\Pi_0$  of  $\Pi$ . We adopt the notation introduced in [PSt15] and, in particular, write  $\mathfrak{z}(\mathfrak{l})_{\text{reg}}$  for the set of all elements  $z$  in the centre  $\mathfrak{z}(\mathfrak{l})$  for which  $\mathfrak{g}_z = \mathfrak{l}$ . This is a nonempty Zariski open subset of  $\mathfrak{z}(\mathfrak{l})$ . Since  $\{0\}$  is a rigid nilpotent orbit in  $\mathfrak{l}$  there exists a unique sheet  $\mathcal{S}$  of  $\mathfrak{g}$  whose open decomposition class coincides with  $\mathcal{D}(\mathfrak{l}, 0) = (\text{Ad } G) \cdot \mathfrak{z}(\mathfrak{l})_{\text{reg}}$ ; see [PSt15, Theorem 2.8]. Furthermore, in view of [PSt15, 2.5] the unique nilpotent orbit contained in  $\mathcal{S}$  coincides with  $\mathcal{O}_0 := \text{Ind}_{\mathfrak{l}}^{\mathfrak{g}} \{0\}$ , the Richardson orbit associated with the Levi subalgebra  $\mathfrak{l}$ . It follows from [PSt15, Theorem 1.4] that the Dynkin label of  $\mathcal{O}_0$  coincides with that of the Richardson orbit in  $\mathfrak{g}_{\mathbb{C}}$  associated with  $\mathfrak{l}_{\mathbb{C}}$ . The latter can be read off from the tables in [dGE09].

Since  $h$  is a semisimple element of  $\mathfrak{g}$  the adjoint  $G$ -orbit  $\mathcal{O}$  of  $h$  is Zariski closed in  $\mathfrak{g}$ . Let  $\mathbb{K}\mathcal{O}$  denote the cone associated with  $\mathcal{O}$ , a  $G$ -stable, Zariski closed, conical subset of  $\mathfrak{g}$ ; see [PSk99, 5.1], for example. It is well known that all irreducible components of  $\mathbb{K}\mathcal{O}$  have dimension equal to that of  $\mathcal{O}$ . Since  $x^{[p]} - x = 0$  for all  $x \in \mathcal{O}$  and the  $p$ -th power map of  $\mathfrak{g}$  is a morphism given by a collection of homogeneous polynomial functions of degree  $p$  on  $\mathfrak{g}$  we have the inclusion  $\mathbb{K}\mathcal{O} \subseteq \mathcal{N}_p(\mathfrak{g})$ . On the other hand, it follows from [PSk99, Lemma 5.1] that  $\mathbb{K}\mathcal{O} = \mathbb{K}^{\times} \mathcal{O} \cup \{0\}$ . Since  $\mathbb{K}^{\times} \mathcal{O} \subseteq \mathcal{D}(\mathfrak{l}, 0) \subset \mathcal{S}$  and  $\mathcal{S}$  contains a unique nilpotent orbit, we now deduce that  $\mathbb{K}\mathcal{O}$  coincides with the Zariski closure of the Richardson orbit  $\mathcal{O}_0$ . As a consequence,  $\mathcal{O}_0 \subseteq \mathcal{N}_p(\mathfrak{g})$ . We let  $e$  be any element of  $\mathcal{O}_0$ . Since  $p$  is a good prime for  $G$  we have that  $\dim \mathfrak{g}_e = \dim G_e = \dim G_h = \dim \mathfrak{g}_h$ . In view of (2) this shows that  $\dim \mathcal{O}_0$  is divisible by  $p(p-1)$ .

(d) Suppose  $G$  is of type  $E_6$  and let  $h, L, \mathfrak{l}, \Psi, \mathcal{O}_0, e$  be as above. If  $p \geq 11$  then  $p(p-1) > 78 = \dim \mathfrak{g}$ . So  $\mathfrak{g}$  cannot contain orbits of dimension divisible by  $p(p-1)$ .

Suppose  $p = 7$ . Then it must be that  $\dim \mathcal{O}_0 = 42$ . There is only one nilpotent orbit of dimension 42 in  $\mathfrak{g}$  and its Dynkin label is  $A_2$ . In view of [dGE09, p. 267] this yields that  $\Psi$  has type  $A_5$ . But then  $\mathfrak{l} = \mathfrak{z}(\mathfrak{l}) \oplus [\mathfrak{l}, \mathfrak{l}]$  and  $\mathfrak{z}(\mathfrak{l}) = \mathbb{K}h$ . Since the restriction of  $\kappa$  to  $\mathfrak{l}$  is non-degenerate and  $\mathfrak{z}(\mathfrak{l}) \subset [\mathfrak{l}, \mathfrak{l}]^{\perp}$ , this forces  $\kappa(h, h) \neq 0$ . Since this contradicts (1), the case  $p = 7$  cannot occur.

Suppose  $p = 5$ . The  $p(p-1) = 20$ . As there are no nilpotent orbits of dimension 20 in  $\mathfrak{g}$  we have that  $\dim \mathcal{O}_0 \in \{40, 60\}$ . The only nilpotent orbit of dimension 40 in  $\mathfrak{g}$  has Dynkin label  $3A_1$ . Since this orbit is rigid by [PSt15, Theorem 1.3], it must be that  $\dim \mathcal{O}_0 = 60$ . There are two nilpotent orbits of dimension 60 in  $\mathfrak{g}$  and their Dynkin labels are  $A_4$  and  $D_4$ . Both orbits are Richardson by [dGE09, p. 267]. However, if  $e \in \mathcal{O}(D_4)$  then  $e^{[5]} \neq 0$  by [McN03, Theorem 35] and [Law95, Table 6]. So it must be that  $e \in \mathcal{O}(A_4)$ . Then [dGE09, p. 267] shows that  $\Psi$  has type  $A_3$ . Let  $\sigma = \sigma(h) \in G_{\mathbb{C}}$  be the element of order  $p$  attached to  $h$  in part (b) and let  $\sigma_{\mathbf{a}} \in T_{\mathbb{C}}$  be a canonical representative of the conjugacy class of  $\sigma$ . Our discussion at the end of part (b) shows that the set  $\{\alpha \in \tilde{\Pi} \mid a_{\alpha} = 0\}$  forms a subdiagram of type  $A_3$  of the extended Dynkin diagram  $\tilde{E}_6$ . Since  $a_0 = 5 - a_1 - 2a_2 - 2a_3 - 3a_4 - 2a_5 - a_6 \geq 0$  and  $a_i \in \mathbb{Z}_{\geq 0}$  for all  $i$ , we essentially have only one option here, namely,

$$(3) \quad (a_0, a_1, a_2, a_3, a_4, a_5, a_6) = (1, 1, 1, 0, 0, 0, 1).$$

The other two options that we have can be obtained from this one by applying a suitable symmetry of  $\tilde{\Pi}$  coming from the group  $\text{Stab}_W(\tilde{\Pi}) \cong \mathbb{Z}/3\mathbb{Z}$ . The uniqueness of  $\sigma_{\mathbf{a}}$  implies that all powers  $\sigma^i$  with  $1 \leq i \leq 4$  are  $G$ -conjugate forcing  $\dim_{\mathbb{C}} \mathfrak{g}_{\mathbb{C}}(\sigma, \zeta^i) = \dim_{\mathbb{C}} \mathfrak{g}_{\mathbb{C}}(\sigma, \zeta^j) = 15$  for all  $1 \leq i, j \leq 4$ . Our discussion at the beginning of part (b) now shows that in characteristic 5 the Lie algebra  $\mathfrak{g}$  admits a unique  $G$ -orbit consisting of  $p$ -balanced toral elements. Statement (i) follows.

(e) Suppose  $G$  is of type  $E_7$ . If  $p \geq 13$  then  $p(p-1) > 133 = \dim \mathfrak{g}$ . So  $\mathfrak{g}$  cannot contain orbits of dimension divisible by  $p(p-1)$ .



If  $p = 11$  then  $\dim \mathcal{O}_0 = 110$  and  $\dim \mathfrak{g}(h, i) = 11$  for all  $i \in \mathbb{F}_p^\times$ . There are two nilpotent orbits of dimension 110 in  $\mathfrak{g}$ , but only one of them is Richardson; see [dGE09, p. 269]. So  $e$  must have Dynkin label  $E_6(a_3)$ . But then  $\Psi$  has type  $A_3A_1^2$  and hence  $h = \lambda t_3 + \mu t_6$  for some  $\lambda, \mu \in \mathbb{F}_p^\times$ . Set

$$\Phi_{3,6}^+ := \{\gamma \in \Phi^+ \mid \nu_3(\gamma) = 1 \text{ and } \nu_6(\gamma) = 0\}.$$

It is straightforward to check that  $|\Phi_{3,6}^+| = 12$ . Since the  $\mathbb{k}$ -span of the root vectors  $e_\gamma$  with  $\gamma \in \Phi_{3,6}^+$  is contained in a single eigenspace  $\mathfrak{g}(h, k)$  with  $k \in \mathbb{F}_p^\times$  we see that  $h$  cannot be  $p$ -balanced. So the case where  $p = 11$  cannot occur.

Suppose  $p = 7$ . Then  $\dim \mathcal{O}_0 \in \{42, 84, 126\}$ . If  $\dim \mathcal{O}_0 = 126$  then  $\mathcal{O}_0 = \mathcal{O}_{\text{reg}}$ , the regular nilpotent orbit of  $\mathfrak{g}$ . Since  $\mathcal{O}_{\text{reg}} \not\subset \mathcal{N}_p(\mathfrak{g})$  by [McN03, Theorem 35] and [Law95, Table 8], this case cannot occur. As  $\mathcal{N}(\mathfrak{g})$  has no orbits of dimension 42 it must be that  $\dim \mathcal{O} = 84$  and  $\dim \mathfrak{g}(h, i) = 14$  for all  $i \in \mathbb{F}_p^\times$ . There are three orbits of this dimension in  $\mathcal{N}(\mathfrak{g})$ , but only two of them are Richardson; see [dGE09, p. 270]. Their Dynkin labels are  $2A_2$  and  $A_2 + 3A_1$ . In the first case  $\Psi$  has type  $D_5A_1$  and hence  $\mathfrak{z}(\mathfrak{l}) = \mathbb{k}h$ . Since  $p \neq 2$  we have a direct sum decomposition  $\mathfrak{l} = \mathfrak{z}(\mathfrak{l}) \oplus [\mathfrak{l}, \mathfrak{l}]$ . Arguing as at the beginning of part (d) it is easy to see that  $\kappa(h, h) \neq 0$ . Since this contradicts (1) we deduce that the present case is impossible. If  $\mathcal{O}_0 = \mathcal{O}(A_2 + 3A_1)$  then  $\Psi$  has type  $A_6$ . In this case we also have the equality  $\mathfrak{z}(\mathfrak{l}) = \mathbb{k}h$ , but cannot argue as before because  $[\mathfrak{l}, \mathfrak{l}] \cong \mathfrak{sl}_7$  contains  $\mathfrak{z}(\mathfrak{l})$ . Instead, we observe that  $h = \lambda t_2$  for some  $\lambda \in \mathbb{F}_p^\times$ . Since the set  $\{\gamma \in \Phi^+ \mid \nu_2(\gamma) = 1\}$  has cardinality  $35 > 14$  we see that  $h$  cannot be  $p$ -balanced in  $\mathfrak{g}$ . Therefore,  $p \neq 7$ .

Suppose  $p = 5$ . Then  $\dim \mathcal{O}_0 \in \{20, 40, 60, 80, 100, 120\}$ . Since there are no orbits of dimension 20, 40, 60 or 80 in  $\mathcal{N}(\mathfrak{g})$  it must be that  $\dim \mathcal{O}_0 = 100$  or  $\dim \mathcal{O}_0 = 120$ . If  $\dim \mathcal{O}_0 = 120$  then  $\mathcal{O}_0 = \mathcal{O}(E_7(a_3))$  or  $\mathcal{O}_0 = \mathcal{O}(E_6)$ . Since neither of these orbits lies in  $\mathcal{N}_p(\mathfrak{g})$  by [McN03, Theorem 35] and [Law95, Table 8], we obtain that  $\dim \mathcal{O}_0 = 100$  and  $\dim \mathfrak{g}(h, i) = 20$  for all  $i \in \mathbb{F}_p^\times$ . There are two orbits of that dimension in  $\mathfrak{g}$  and their Dynkin labels are  $A_4$  and  $A_3 + A_2 + A_1$ . If  $\mathcal{O}_0 = \mathcal{O}(A_3 + A_2 + A_1)$  then [dGE09, p. 269] yields that  $\Psi$  has type  $A_4A_2$  and  $\mathfrak{z}(\mathfrak{l}) = \mathbb{k}t_5$ . As a consequence,  $h = \lambda t_5$  for some  $\lambda \in \mathbb{F}_p^\times$ . Since the set  $\{\gamma \in \Phi^+ \mid \nu_5(\gamma) = 1\}$  has cardinality  $30 > 15$ , the element  $h$  cannot be  $p$ -balanced in the present case. Now suppose  $\mathcal{O}_0 = \mathcal{O}(A_4)$ . Then [dGE09, p. 267] shows that  $\Phi$  is of type  $D_4A_1$ . Let  $\sigma = \sigma(h) \in G_{\mathbb{C}}$  be the element of order  $p$  attached to  $h$  in part (b) and let  $\sigma_{\mathbf{a}} \in T_{\mathbb{C}}$  be a canonical representative of the conjugacy class of  $\sigma$ . Our discussion at the end of part (b) shows that the set  $\{\alpha \in \tilde{\Pi} \mid a_\alpha = 0\}$  forms a subdiagram of type  $D_4A_1$  on the extended Dynkin diagram  $\tilde{E}_7$ . Since  $a_0 = 5 - 2a_1 - 2a_2 - 3a_3 - 4a_4 - 3a_5 - 2a_6 - a_7 \geq 0$  and  $a_i \in \mathbb{Z}_{\geq 0}$  for all  $i$ , we essentially have only one option here, namely,

$$(4) \quad (a_0, a_1, a_2, a_3, a_4, a_5, a_6, a_7) = (1, 1, 0, 0, 0, 0, 1, 0).$$

The other option that we have can be obtained by applying a suitable symmetry of  $\tilde{\Pi}$  coming from the group  $\text{Stab}_W(\tilde{\Pi}) \cong \mathbb{Z}/2\mathbb{Z}$ . As at the end of part (d), the uniqueness of  $\sigma_{\mathbf{a}}$  implies that all powers  $\sigma^i$  with  $1 \leq i \leq 4$  are  $G$ -conjugate. Then  $\dim_{\mathbb{C}} \mathfrak{g}_{\mathbb{C}}(\sigma, \zeta^i) = \dim_{\mathbb{C}} \mathfrak{g}_{\mathbb{C}}(\sigma, \zeta^j) = 20$  for all  $1 \leq i, j \leq 4$  (of course, one can also check this by a direct computation). Our discussion at the beginning of part (b) now shows that in characteristic 5 the Lie algebra  $\mathfrak{g}$  admits a unique  $G$ -orbit consisting of  $p$ -balanced toral elements. This proves statement (ii).

(f) Suppose  $G$  is of type  $E_8$ . If  $p \geq 17$  then  $p(p-1) > 248$ , hence  $\mathfrak{g}$  does not have orbits of dimension divisible by  $p(p-1)$ . If  $p = 13$  then it must be that  $\dim \mathcal{O}_0 = 156$ . There is only one orbit of dimension 156 in  $\mathcal{N}(\mathfrak{g})$  and its Dynkin label is  $2A_2$ . Then [dGE09, p. 275] yields that  $\Psi$  has type  $D_7$  and  $\mathfrak{z}(\mathfrak{l}) = \mathbb{k}t_1$ . Since  $\mathfrak{l} = \mathfrak{z}(\mathfrak{l}) \oplus [\mathfrak{l}, \mathfrak{l}]$  we then have  $\kappa(h, h) \neq 0$  contrary to (1). Therefore,  $p \leq 11$ .

Suppose  $p = 11$ . Then  $\dim \mathcal{O}_0 = 220$  and  $\dim \mathfrak{g}(h, i) = 22$  for all  $i \in \mathbb{F}_p^\times$  because  $\mathcal{N}(\mathfrak{g})$  has no orbits of dimension 110. There are two orbits of dimension 220 in  $\mathcal{N}(\mathfrak{g})$  and their Dynkin labels are  $E_7(a_3)$  and  $E_8(b_6)$ . Both orbits are Richardson and it follows from [dGE09, p. 272] that in the second case  $\Psi$  has type  $A_3A_2A_1$  and  $h = \lambda t_4 + \mu t_8$  for some  $\lambda, \mu \in \mathbb{F}_p^\times$ . Set

$$\Phi_{4,8}^+ := \{\gamma \in \Phi^+ \mid \nu_4(\gamma) = 1 \text{ and } \nu_8(\gamma) = 0\}.$$

It is easy to see that  $|\Phi_{4,8}^+| = 24 > 22$ . Since the  $\mathbb{k}$ -span of the root vectors  $e_\gamma$  with  $\gamma \in \Phi_{4,8}^+$  is contained in a single eigenspace of  $\text{ad } h$  we deduce that the case where  $\mathcal{O}_0 = \mathcal{O}(E_8(b_6))$  cannot occur. If  $\mathcal{O}_0 = \mathcal{O}(E_7(a_3))$  then [dGE09, p. 272] shows that  $\Psi$  has type  $A_4$ . Let  $\sigma = \sigma(h) \in G_{\mathbb{C}}$  be the element of order  $p$  attached to  $h$  in part (b) and let  $\sigma_{\mathbf{a}} \in T_{\mathbb{C}}$  be a canonical representative of the conjugacy class of  $\sigma$ . We know from part (b) that the set  $\{\alpha \in \tilde{\Pi} \mid a_\alpha = 0\}$  forms a subdiagram of type  $A_4$  of the extended Dynkin diagram  $\tilde{E}_8$ . Since  $a_0 = 11 - 2a_1 - 3a_2 - 4a_3 - 6a_4 - 5a_5 - 4a_6 - 3a_7 - 2a_8 \geq 0$  and  $a_i \in \mathbb{Z}_{\geq 0}$  for all  $i$ , we have only one option here, namely,

$$(5) \quad (a_0, a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8) = (1, 1, 1, 0, 0, 0, 0, 1, 1).$$

The uniqueness of  $\sigma_{\mathbf{a}}$  implies that all powers  $\sigma^i$  with  $1 \leq i \leq 10$  are  $G$ -conjugate implying that  $\dim_{\mathbb{C}} \mathfrak{g}_{\mathbb{C}}(\sigma, \zeta^i) = \dim_{\mathbb{C}} \mathfrak{g}_{\mathbb{C}}(\sigma, \zeta^j) = 22$  for all  $1 \leq i, j \leq 10$ . Arguing as before we now conclude that in characteristic 11 the Lie algebra  $\mathfrak{g}$  admits a unique  $G$ -orbit consisting of  $p$ -balanced toral elements. This proves (v).

Suppose  $p = 7$ . Since  $\mathcal{N}(\mathfrak{g})$  has no orbits of dimension 42, 84 and 126, either  $\dim \mathcal{O}_0 = 168$  or  $\dim \mathcal{O}_0 = 210$ . There are two orbits of dimension 168 in  $\mathcal{N}(\mathfrak{g})$ , but one of them is rigid by [PSt15, Theorem 1.3]. So if  $\dim \mathcal{O}_0 = 168$  then  $\dim \mathfrak{g}(h, i) = 28$  for all  $i \in \mathbb{F}_p^\times$  and  $\mathcal{O}_0 = \mathcal{O}(D_4)$ . In that case [dGE09, p. 275] yields that  $\Psi$  has type  $E_6$  and  $h = \lambda t_7 + \mu t_8$  for some  $\lambda, \mu \in \mathbb{F}_p^\times$ . Let  $\sigma = \sigma(h) \in G_{\mathbb{C}}$  be the element of order  $p$  attached to  $h$  in part (b) and let  $\sigma_{\mathbf{a}} \in T_{\mathbb{C}}$  be a canonical representative of the conjugacy class of  $\sigma$ . From part (b) we know that the set  $\{\alpha \in \tilde{\Pi} \mid a_\alpha = 0\}$  forms a subdiagram of type  $E_6$  of the extended Dynkin diagram  $\tilde{E}_8$ . Since  $a_0 = 7 - 2a_1 - 3a_2 - 4a_3 - 6a_4 - 5a_5 - 4a_6 - 3a_7 - 2a_8 \geq 0$  and  $a_i \in \mathbb{Z}_{\geq 0}$  for all  $i$ , we have only one option here, namely,

$$(6) \quad (a_0, a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8) = (2, 0, 0, 0, 0, 0, 0, 1, 1).$$

The uniqueness of  $\sigma_{\mathbf{a}}$  implies that all  $\sigma^i$  with  $1 \leq i \leq 6$  are  $G$ -conjugate forcing  $\dim_{\mathbb{C}} \mathfrak{g}_{\mathbb{C}}(\sigma, \zeta^i) = \dim_{\mathbb{C}} \mathfrak{g}_{\mathbb{C}}(\sigma, \zeta^j) = 28$  for all  $1 \leq i, j \leq 6$ . Therefore, in the present case  $h$  is a  $p$ -balanced toral element of  $\mathfrak{g}$ . Statement (iv) follows.

It remains to consider the case where  $p = 7$  and  $\dim \mathcal{O}_0 = 210$ . There are two orbits of that dimension in  $\mathcal{N}(\mathfrak{g})$  and their Dynkin labels are  $A_6$  and  $D_4(a_1)$ . Both orbits are Richardson by [dGE, p. 273]. As  $\mathcal{O}_0 = \mathcal{O}(D_4(a_1)) \not\subset \mathcal{N}_p(\mathfrak{g})$  by [McN03, Theorem 35] and [Law95, Table 9] it must be that  $\mathcal{O}_0 = \mathcal{O}(A_6)$ . Then [dGE09, p. 273] in conjunction with [PSt15, Theorem 1.4] shows that  $\Psi$  is of type  $D_4A_2$  and  $h = \lambda t_1 + \mu t_6$  for some  $\lambda, \mu \in \mathbb{F}_p^\times$ . Let  $\sigma_{\mathbf{a}} \in T_{\mathbb{C}}$  be a canonical representative of the conjugacy class of  $\sigma(h)$ . By part (b), the set  $\{\alpha \in \tilde{\Pi} \mid a_\alpha = 0\}$  forms a subdiagram of type  $D_4A_2$  of the extended Dynkin diagram  $\tilde{E}_8$ . Since  $a_0 = 7 - 2a_1 - 3a_2 - 4a_3 - 6a_4 - 5a_5 - 4a_6 - 3a_7 - 2a_8 \geq 0$  and  $a_i \in \mathbb{Z}_{\geq 0}$  for all  $i$ , the only option we have here is

$$(7) \quad (a_0, a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8) = (1, 1, 0, 0, 0, 0, 1, 0, 0).$$

Once again the uniqueness of  $\sigma_{\mathbf{a}}$  implies that all powers  $\sigma^i$  with  $1 \leq i \leq 6$  are  $G$ -conjugate. Hence  $\dim_{\mathbb{C}} \mathfrak{g}_{\mathbb{C}}(\sigma, \zeta^i) = \dim \mathfrak{g}(h, i) = 35$  for all  $i \in \mathbb{F}_p^\times$ . Therefore,  $h$  is a  $p$ -balanced toral element of  $\mathfrak{g}$ , proving (v).

(g) Suppose  $G$  is of type  $F_4$ . If  $p > 7$  then  $p(p-1) > 52 = \dim \mathfrak{g}$ , hence  $\mathfrak{g}$  has no orbits of dimension divisible by  $p(p-1)$ . If  $p = 7$  then it must be that  $\dim \mathcal{O}_0 = 42$  and  $\dim \mathfrak{g}(h, i) = 7$  for all  $i \in \mathbb{F}_p^\times$ . There are two orbits of dimension 42 in  $\mathcal{N}(\mathfrak{g})$  and their Dynkin labels are  $B_3$  and  $C_3$ . By [dGE09, p. 276] (which is applicable in view of [PSt15, Theorem 1.4]), both orbits are Richardson and  $\Psi$  has type  $A_2$  in both cases (in the  $B_3$ -case  $\Psi$  consists of short roots whereas in the  $C_3$ -case all roots in  $\Psi$  are long). In both cases there exists a unique collection  $\mathbf{a} = (a_0, a_1, a_2, a_3, a_4)$  with  $a_i \in \mathbb{Z}_{\geq 0}$  and  $a_0 = 7 - 2a_1 - 3a_2 - 4a_3 - 2a_4$  such that the set  $\{\alpha \in \tilde{\Pi} \mid a_\alpha = 0\}$  forms a basis of  $\Psi$ . Specifically, in the  $B_3$ -case we take  $\mathbf{a} = (2, 1, 1, 0, 0)$  whilst in the  $C_3$ -case we take  $\mathbf{a} = (1, 0, 0, 1, 1)$  (these are the only options available). Arguing as before it is now easy to observe that each of these two cases gives rise to a unique  $p$ -balanced conjugacy class in  $\mathfrak{g}$ . This proves (vii).

Suppose  $p = 5$ . Since there are no orbits of dimension 20 in  $\mathfrak{g}$  and there exists only one orbit of dimension 40 in  $\mathcal{N}(\mathfrak{g})$  it must be that  $\mathcal{O}_0 = \mathcal{O}(F_4(a_3))$ . Thanks to [dGE09, p. 276] this yields that  $[\mathfrak{l}, \mathfrak{l}]$  has type  $A_1 A_2$  and  $h \in \mathbb{k}^\times t_2$  spans the centre of  $\mathfrak{l}$ . Since  $p = 5$  we have that  $\mathfrak{l} = \mathfrak{z}(\mathfrak{l}) \oplus [\mathfrak{l}, \mathfrak{l}]$  this forces  $\kappa(h, h) \neq 0$  contrary to (1). Therefore, the case  $p = 5$  does not produce any  $p$ -balanced elements.

If  $G$  is of type  $G_2$  then  $p(p-1) > \dim \mathfrak{g} = 14$  for any good prime  $p$ . This means that  $\mathfrak{g}$  does not admit any  $p$ -balanced toral elements. This completes the proof of the proposition.  $\square$

2.8. In the final stages of the proof of Theorem 1.1 we shall require some information on the gradings of non-restricted Hamiltonian algebras  $H(2; \underline{n}; \Phi)^{(1)}$ .

Given  $m \in \mathbb{Z}_{\geq 1}$  we denote by  $O((m))$  the full divided power algebra in  $m$  variables over  $\mathbb{k}$ . Its elements are the infinite series  $\sum_{\mathbf{r}} \lambda_{\mathbf{r}} x^{\mathbf{r}}$  with  $\lambda_{\mathbf{r}} \in \mathbb{k}$ . Here  $x^{\mathbf{r}} = x_1^{(r_1)} \cdots x_m^{(r_m)}$ , where  $r_i \in \mathbb{Z}_{\geq 0}$ , and the product is induced by the rule

$$x^{\mathbf{a}} \cdot x^{\mathbf{b}} = \left( \prod_{i=1}^m \binom{a_i + b_i}{a_i} \right) \cdot x^{\mathbf{a} + \mathbf{b}}.$$

The maximal ideal  $\mathfrak{m}$  of the ring  $O((m))$  is equipped with a system of divided powers  $f \mapsto f^{(k)} \in O((m))$  where  $k \in \mathbb{Z}_{\geq 0}$ ; see [Str04, 2.1] for more detail. Given  $f \in \mathfrak{m}$  we set

$$\exp(f) := \sum_{k \geq 0} f^{(k)}.$$

This element is well-defined in the linearly compact ring  $O((m))$  and using the axioms of divided powers it is straightforward to check that  $\exp(f)^{-1} = \exp(-f)$ .

A continuous derivation  $D$  of the topological algebra  $O((m))$  is called *special* if  $D(f^{(k)}) = f^{(k-1)} D(f)$  for all  $f \in \mathfrak{m}$  and  $k > 0$ . The special derivations of  $O((m))$  form a Lie subalgebra of  $\text{Der}(O((m)))$  denoted  $W((m))$ . It is well known that this algebra is a free  $O((m))$ -module of rank  $m$  with a free basis consisting of the special partial derivatives  $\partial_1, \dots, \partial_m$ . Recall that  $\partial_i(x^{(r)}) = x^{(r - \epsilon_i)}$  where  $\epsilon_i = (\delta_{i,1}, \dots, \delta_{i,m})$  and  $1 \leq i \leq m$ .

If  $\underline{n} = (n_1, \dots, n_m) \in \mathbb{Z}_{\geq 1}^m$  then the  $\mathbb{k}$ -span  $O(m; \underline{n})$  of all  $x^{\mathbf{r}}$  with  $0 \leq r_i \leq p^{n_i} - 1$  is a subalgebra of dimension  $p^{n_1 + \dots + n_m}$  in  $O((m))$  invariant under all special partial derivatives  $\partial_i$  with  $1 \leq i \leq m$ . The *general Cartan type Lie algebra*  $W(m; \underline{n}) = W((m)) \cap \text{Der}(O(m; \underline{n}))$  is a free  $O(m; \underline{n})$ -module with basis  $\partial_1, \dots, \partial_m$ .

We denote by  $\Omega((m)) = \bigoplus_{i=0}^m \Omega^i((m))$  the module of Kähler differentials over  $O((m))$ . It is known that any Hamiltonian algebra  $H(2; \underline{n}; \Phi)^{(1)} \subset W(2; \underline{n})$  has Cartan type  $S$  and stabilises the volume form  $J(\Phi)\omega_S \in \Omega^2((2))$  where  $\omega_S = dx_1 \wedge dx_2$ . Here  $J(\Phi) \in O((2))^\times$  is the Jacobian of the admissible

automorphism  $\Phi \in \text{Aut}_c O((2))$ ; see [Str04, p. 301]. Moreover, thanks to [Str04, Theorem 6.3.8] we may assume that  $\Phi = \Phi(\tau)$  or  $\Phi = \Phi(l)$  where  $\tau = \tau(\underline{n}) = (p^{n_1} - 1, p^{n_2} - 1)$  and  $l = 1, 2$ . Recall that  $J(\Phi(\tau)) = 1 + x^{\tau(\underline{n})}$  and  $J(\Phi(l)) = \exp(x_l^{(p^{n_l})})$ ; see [Str04, p. 309].

**Lemma 2.8.** *Let  $S = H(2; \underline{n}; \Phi)^{(1)}$  where  $\Phi$  is one of  $\Phi(\tau)$ ,  $\Phi(1)$ ,  $\Phi(2)$ . Then all maximal tori of the algebraic group  $\text{Aut}(S)$  are 1-dimensional.*

*Proof.* Let  $\mathbf{G} = \text{Aut}(S)$  and denote by  $\tilde{\mathbf{G}}$  the normaliser of  $O(2; \underline{n})$  in  $\text{Aut}_c O((2))$ . It follows from [Skr91, Theorem 10.8] that the group  $\mathbf{G}$  consists of all  $g \in \tilde{\mathbf{G}}$  such that  $g(J(\Phi)\omega_S) \in \mathbb{k}^\times J(\Phi)\omega_S$  (see also [Str04, Theorems 7.3.2]). Any element  $g \in \tilde{\mathbf{G}}$  is uniquely determined by its effect on the divided power generators  $x_1, x_2 \in O((2))$  and  $g(x_1), g(x_2) \in O(2; \underline{n})$ . From this it is immediate that  $\tilde{\mathbf{G}}/R_u(\tilde{\mathbf{G}})$  is isomorphic to a parabolic subgroup of  $\text{GL}_2(\mathbb{k})$ . It follows that all maximal tori of  $\tilde{\mathbf{G}}$  have dimension 2. One such torus,  $T_0$ , consists of all automorphisms  $g(t_1, t_2)$  with  $t_1, t_2 \in \mathbb{k}^\times$  such that  $(g(t_1, t_2))(x_i) = t_i x_i$  for  $i = 1, 2$ .

Suppose  $\mathbf{G}$  contains a 2-dimensional torus. Then it is a maximal torus of  $\tilde{\mathbf{G}}$  and hence is  $\tilde{\mathbf{G}}$ -conjugate to  $T_0$ . Replacing  $\Phi \in \text{Aut}_c O((2))$  by a  $g\Phi g^{-1}$  for a suitable  $g \in \tilde{\mathbf{G}}$  we may assume  $T_0 \subset \mathbf{G}$ . Then  $J(\Phi)dx_1 \wedge dx_2$  is a weight vector for the action of  $T_0$  on  $\Omega^2((2))$ . Note that  $O((2))$  decomposes into an infinite direct sum of weight spaces for  $T_0$ . Since  $(g(t_1, t_2))(x^{\mathbf{a}}) = t_1^{a_1} t_2^{a_2} \cdot x^{\mathbf{a}}$ , all  $T_0$ -weight spaces of  $O((2))$  are 1-dimensional and spanned by monomials  $x^{\mathbf{a}}$ . As a consequence the  $T_0$ -weight spaces of  $\Omega^2((2))$  have the form  $\mathbb{k}x^{\mathbf{a}}dx_1 \wedge dx_2$  for  $\mathbf{a} \in \mathbb{Z}_{\geq 0}^2$ . Since  $J(\Phi)$  is invertible in  $O((2))$  and  $J(\Phi)dx_1 \wedge dx_2 = \lambda x^{\mathbf{a}}dx_1 \wedge dx_2$  for some  $\lambda \in \mathbb{k}^\times$  and  $\mathbf{a} \in \mathbb{Z}_{\geq 0}^2$ , it must be that  $\mathbf{a} = (0, 0)$ . But then  $H(2; \underline{n}, \Phi)^{(1)} \subseteq H(2; \underline{n})$ . The simplicity of  $H(2; \underline{n}, \Phi)^{(1)}$  now forces  $H(2; \underline{n}, \Phi)^{(1)} \subseteq H(2; \underline{n})^{(2)}$ . Since  $p^{n_1+n_2} - 2 = \dim H(2; \underline{n})^{(2)} \leq \dim H(2; \underline{n}, \Phi)^{(1)} \leq p^{n_1+n_2} - 1$  this is impossible. This contradiction shows that the maximal tori of  $\mathbf{G}$  are at most 1-dimensional.

On the other hand, using the above expressions for  $J(\Phi)$  it is straightforward, in each case, to produce a 1-dimensional subtorus  $T_\Phi$  of  $T_0$  contained in  $\mathbf{G}$ . If  $\Phi = \Phi(\tau)$  we take for  $T_\Phi$  the identity component of  $\{g(t_1, t_2) \in T_0 \mid t_1^{p^{n_1}-1} t_2^{p^{n_2}-1} = 1\}$  and if  $\Phi = \Phi(l)$ , where  $l \in \{1, 2\}$ , we take  $T_\Phi = \{g(t_1, t_2) \in T_0 \mid t_l = 1\}$ . This completes the proof.  $\square$

### 3. PROOF OF THE MAIN THEOREM

3.1. Our proof of Theorem 1.1 will make essential use of Weisfeiler's theorem; see [Weis78] and [Str04, Theorems 3.5.6, 3.5.7 ad 3.5.8]. Let  $O(m; \underline{n})$  and  $W(m, \underline{n})$  be as in (2.8). We shall often encounter the case where  $\underline{n} = \underline{1}$ , that is  $n_i = 1$  for all  $i$ . In this case all derivations of  $O(m; \underline{n})$  are special and  $W(m; \underline{1})$  is the full derivation algebra of  $O(m; \underline{1}) \cong \mathbb{k}[X_1, \dots, X_m]/(X_1^p, \dots, X_m^p)$ . For  $1 \leq i \leq m$  we let  $x_i$  stand for the divided-power generator  $x^{(\epsilon_i)}$  of  $O(m; \underline{n})$  and denote by  $O(x_{i_1}, \dots, x_{i_s})$  the divided power subalgebra of  $O(m; \underline{n})$  generated by  $x_{i_1}, \dots, x_{i_s}$  (as a  $\mathbb{k}$ -algebra  $O(x_{i_1}, \dots, x_{i_s})$  is isomorphic to a truncated polynomial ring in  $n_{i_1} + \dots + n_{i_s}$  variables). A Lie subalgebra  $\mathcal{D}$  of  $W(m; \underline{n}) = \sum_{i=1}^m O(m; \underline{n}) \partial_i$  is called *transitive* if  $O(m; \underline{n})$  does not contain nonzero proper  $\mathcal{D}$ -invariant ideals.

3.2. Given a non-empty subset  $X \subseteq \mathfrak{g}$  and  $k \in \mathbb{Z}_{\geq 2}$  we let  $X^k$  denote the linear span of all elements  $((\text{ad } x_1) \circ \dots \circ (\text{ad } x_{k-1}))(x_k)$  with  $x_i \in X$ . Put  $M_{(0)} := M$  and choose an  $(\text{ad } M_{(0)})$ -stable subspace  $M_{(-1)}$  in  $\mathfrak{g}$  such that  $M_{(0)} \subsetneq M_{(-1)}$  and  $M_{(-1)}/M_{(0)}$  is an irreducible  $M_{(0)}$ -module. Given  $i \in \mathbb{Z}_{\geq 2}$  we define a subspace  $M_{(-i)} \subseteq \mathfrak{g}$  recursively by setting  $M_{(-i)} = M_{(-1)}^i + M_{(-i+1)}$ . Since  $M$  is a maximal

subalgebra of  $\mathfrak{g}$  there exists a positive integer  $q$  such that  $\mathfrak{g} = M_{(-q)} \supsetneq M_{(-q+1)}$ . Set  $M_{(1)} := \{x \in M_{(0)} \mid [x, M_{(-1)}] \subseteq M_{(0)}\}$  and for  $i \geq 2$  define  $M_{(i)} := \{x \in M_{(i-1)} \mid [x, M_{(-1)}] \subseteq M_{(i-1)}\}$ . Since  $\mathfrak{g}$  is a simple Lie algebra there is a non-negative integer  $r$  such that  $0 = M_{(r+1)} \subsetneq M_{(r)}$ . It is well known that the chain of subspaces

$$\mathfrak{g} = M_{(-q)} \supset \cdots \supset M_{(0)} \supseteq \cdots \supseteq M_{(r)} \supset 0$$

is a Lie algebra filtration, that is,  $[M_{(i)}, M_{(j)}] \subseteq M_{(i+j)}$  for all  $i, j \in \mathbb{Z}$ . By maximality,  $M$  is a restricted subalgebra of  $\mathfrak{g}$  and can see by induction on  $k$  that  $M_{(k)} = \{x \in M_{(0)} \mid (\text{ad } M_{(-1)}^k)(x) \subseteq M_{(0)}\}$  for all  $k > 0$ . From this it is immediate that  $M_{(k)}^{[p]} \subseteq M_{(pk)}$  for such  $k$ . The decreasing filtration of  $\mathfrak{g}$  thus obtained is referred to as the *Weisfeiler filtration* associated with the pair  $(M_{(-1)}, M_{(0)})$ .

Since  $M_{(-1)}/M_{(0)}$  is an irreducible module over  $M_{(0)} = M$  we have that  $\text{nil}(M) \subseteq M_{(1)}$ . If  $x \in M_{(1)}$  then  $(\text{ad } x)^{q+r+1}(M_{(-q)}) = 0$ . So the ideal  $M_{(1)}$  of  $M$  consists of nilpotent elements of  $\mathfrak{g}$ . This shows that  $M_{(1)} = \text{nil}(M)$ . Since we are assuming that  $\text{nil}(M) \neq 0$ , Weisfeiler's filtration does not collapse completely in our case. In other words,  $r > 0$ .

We denote by  $\mathcal{G}$  the corresponding graded Lie algebra

$$\mathcal{G} = \text{gr}(\mathfrak{g}) := \bigoplus_{i=-q}^r \mathcal{G}_i, \quad \mathcal{G}_i = \text{gr}_i(\mathfrak{g}) = M_{(i)}/M_{(i+1)},$$

a graded vector space whose Lie product is induced by that of  $\mathfrak{g}$ . We can give  $\mathcal{G}_{\geq 0} := \bigoplus_{i \geq 0} \mathcal{G}_i$  a restricted Lie algebra structure by setting  $(\text{gr}_i(x))^{[p]} := \text{gr}_{pi}(x^{[p]})$  for all  $x \in M_{(i)}$  where  $i \geq 0$ . By construction,  $\mathcal{G}_0 \cong M/\text{nil}(M)$  as Lie algebras,  $\mathcal{G}_{-1}$  is an irreducible  $\mathcal{G}_0$ -module, and the Lie subalgebra  $\mathcal{G}_- := \bigoplus_{i < 0} \mathcal{G}_i$  is generated by  $\mathcal{G}_{-1}$ .

3.3. Let  $N(\mathcal{G}) = \bigoplus_{i < 0} N_i(\mathcal{G})$  be the largest graded ideal of  $\mathcal{G}$  contained in  $\mathcal{G}_-$  and put  $\bar{\mathcal{G}} = \mathcal{G}/N(\mathcal{G})$  so that  $\bar{\mathcal{G}} = \bigoplus_{i=-q}^r \bar{\mathcal{G}}_i$  where  $\bar{\mathcal{G}}_i = \mathcal{G}_i/N_i(\mathcal{G})$ . Since  $N(\mathcal{G}) \cap \mathcal{G}_{-1} = 0$  by the irreducibility of  $\mathcal{G}_{-1}$  we have that  $N(\mathcal{G}) \subseteq \bigoplus_{i \leq -2} \mathcal{G}_i$  (this means that  $\bigoplus_{i \geq -1} \mathcal{G}_i \cong \bigoplus_{i \geq -1} \bar{\mathcal{G}}_i$  as graded vector spaces).

Since  $r > 0$ , the first part of Weisfeiler's theorem says that  $N(\mathcal{G})$  coincides with the radical of  $\mathcal{G}$  and the semisimple Lie algebra  $\bar{\mathcal{G}}$  has a unique minimal ideal which contains  $\bar{\mathcal{G}}_- := \bigoplus_{i \leq -1} \bar{\mathcal{G}}_i$ . This ideal, denoted  $A(\bar{\mathcal{G}})$ , is isomorphic to  $S \otimes O(m; \underline{n})$  for some simple Lie algebra  $S$  and some divided power algebra  $O(m; \underline{n})$ . The restricted Lie algebra structure of  $\mathcal{G}_{\geq 0}$  induces that on  $\bar{\mathcal{G}}_{\geq 0} := \bigoplus_{i \geq 0} \bar{\mathcal{G}}_i$ .

Thanks to Block's theorem, the adjoint action of  $\bar{\mathcal{G}}$  on  $A(\bar{\mathcal{G}})$  gives rise to inclusions

$$(8) \quad S \otimes O(m; \underline{n}) \subset \bar{\mathcal{G}} \subseteq ((\text{Der}(S) \otimes O(m; \underline{n})) \rtimes (\text{Id}_S \otimes W(m; \underline{n})))$$

and the canonical projection

$$\pi: ((\text{Der}(S) \otimes O(m; \underline{n})) \rtimes (\text{Id}_S \otimes W(m; \underline{n}))) \twoheadrightarrow W(m; \underline{n})$$

maps  $\bar{\mathcal{G}}$  onto a transitive subalgebra of  $W(m; \underline{n})$ ; see [Str04, Theorem 3.3.5], for example. Since  $\bar{\mathcal{G}}$  is semisimple (hence centreless) and acts faithfully on  $A(\bar{\mathcal{G}})$ , the restricted Lie algebra  $\bar{\mathcal{G}} \cong \text{ad}_{A(\bar{\mathcal{G}})}(\bar{\mathcal{G}})$  identifies naturally with a restricted subalgebra of  $\text{Der}(A(\bar{\mathcal{G}}))$ .

3.4. The uniqueness of  $A(\bar{\mathcal{G}})$  implies that it is a graded ideal of  $\bar{\mathcal{G}}$ . The grading of  $A(\bar{\mathcal{G}})$  is completely determined by the second part Weisfeiler's theorem which states that only one of the following two cases can occur:



*Degenerate Case.* If  $A(\bar{\mathcal{G}})$  intersects trivially with  $\bar{\mathcal{G}}_+ := \bigoplus_{i>0} \bar{\mathcal{G}}_i$  then necessarily  $m \geq 1$  and the grading of  $A(\bar{\mathcal{G}})$  is induced by that of the associative algebra  $O(m; \underline{n})$ . The latter is given by fixing a positive integer  $s \leq m$  and assigning to the divided power generators  $x_1, \dots, x_s$  and  $x_{s+1}, \dots, x_m$  degrees  $-1$  and  $0$ , respectively. More precisely, for  $i \leq 0$  the graded component of  $A(\bar{\mathcal{G}})$  has the form  $A_i(\bar{\mathcal{G}}) = S \otimes O(x_1, \dots, x_s)[i; s]$  where  $O(x_1, \dots, x_s)[i; s]$  is the linear span of all monomials  $\prod_{i=1}^m x_i^{(a_i)}$  with  $0 \leq a_i \leq p^{n_i} - 1$  and  $a_1 + \dots + a_s = -i$ . Moreover, in this case  $[[\mathcal{G}_{-1}, \mathcal{G}_1], \mathcal{G}_1] = 0$  and  $\mathcal{G}_i = 0$  for all  $i \geq 2$ . Finally,  $\mathcal{G}_1$  identifies with a nonzero subspace of  $\sum_{i=1}^s O(x_{s+1}, \dots, x_m) \partial_i$  through an embedding described in (8) and

$$S \otimes O(x_{s+1}, \dots, x_m) \subset \mathcal{G}_0 \subseteq (\text{Der}(S) \otimes O(x_{s+1}, \dots, x_m)) \rtimes (\text{Id}_S \otimes W(m, \underline{n})[0; s])$$

where  $W(m; \underline{n})[0; s] = \sum_{i>s} O(x_{s+1}, \dots, x_m) \partial_i \oplus \sum_{i,j \leq s} O(x_{s+1}, \dots, x_m)(x_i \partial_j)$ .

*Non-degenerate Case.* The simple Lie algebra  $S$  is  $\mathbb{Z}$ -graded in such a way that  $S_1 \neq 0$  and  $A_i(\bar{\mathcal{G}}) = S_i \otimes O(m; \underline{n})$  for all  $i \in \mathbb{Z}$ . The grading of  $S$  induces that on  $\text{Der}(S)$  and  $\mathcal{G}_0$  is sandwiched between  $S_0 \otimes O(m; \underline{n})$  and  $(\text{Der}_0(S) \otimes O(m; \underline{n})) \rtimes (\text{Id}_S \otimes W(m; \underline{n}))$ . Since  $\mathcal{G}_1 = S_{-1} \otimes (m; \underline{n})$  is a faithful  $\mathcal{G}_0$ -module and  $S_0 \otimes 1 \subseteq S_0 \otimes O(m; \underline{n}) \subseteq \mathcal{G}_0$ , the Lie subalgebra  $S_0$  of  $S$  acts faithfully on  $S_{-1}$ . Although  $S_{-1}$  does not have to be irreducible over  $S_0$ , in general, it follows from [Str04, Theorem 3.5.7(6)] that this module is semisimple and isogenic when  $\text{Der}(S) = \text{ad } S$ .

Let  $M$  be a counterexample to Theorem 1.1. By Lemma 2.4,  $M$  is not regular in  $\mathfrak{g}$ . We set  $M := M_{(0)}$ , choose a subspace  $M_{(-1)}$  as in (3.2), and consider the Weisfeiler filtration of  $\mathfrak{g}$  associated with the pair  $(M_{(-1)}, M_{(0)})$ . Write  $\mathcal{G}$  for the corresponding graded Lie algebra, and let  $N(\mathcal{G})$ ,  $\bar{\mathcal{G}}$  and  $A(\bar{\mathcal{G}}) = S \otimes O(m; \underline{n})$  be as in (3.3).

3.5. In the next three subsections we assume that the degenerate case of Weisfeiler's theorem holds for  $\bar{\mathcal{G}}$ . In particular, this means that  $m \geq 1$ . If  $m = 1$  then necessarily  $s = 1$ . As a consequence,  $\mathcal{G}_1 = \mathbb{k} \partial_1$  is 1-dimensional. Since  $\mathcal{G}_2 = 0$  by (3.4), it must be that  $M_{(2)} = 0$ . So  $\text{nil}(M) = M_{(1)}$  is isomorphic to  $\mathcal{G}_1$ . But then  $\text{nil}(M) = \mathbb{k}e$  for some nonzero nilpotent element  $e \in \mathfrak{g}$  and  $M \subseteq \mathfrak{n}_{\mathfrak{g}}(\mathbb{k}e)$ . This, however, contradicts our assumption on  $M$  because  $\mathfrak{n}_{\mathfrak{g}}(\mathbb{k}e)$  is contained in a proper parabolic subalgebra of  $\mathfrak{g}$  (see the proof of Corollary 2.3 for detail).

Now suppose  $m \geq 2$ . Since  $S$  is a simple Lie algebra and  $\dim O(m; \underline{n}) \geq p^m$ , we have the inequality  $(\dim \mathfrak{g})/p^m > \dim S \geq 3$ . Since  $\mathfrak{g}$  is exceptional and  $p$  is a good prime for  $\mathfrak{g}$ , this forces  $m = 2$ ,  $\underline{n} = (1, 1)$ , and rules out the cases where  $\mathfrak{g}$  is of type  $G_2$  or  $F_4$ . Furthermore, if  $\mathfrak{g}$  is of type  $E_6$  then  $p = 5$  and  $S \cong \mathfrak{sl}_2$ , if  $\mathfrak{g}$  is of type  $E_7$  then  $p = 5$  and  $S$  is either  $\mathfrak{sl}_2$  or  $W(1; \underline{1})$ , and if  $\mathfrak{g}$  is of type  $E_8$  then  $p = 7$  and  $S \cong \mathfrak{sl}_2$ . In any event, all derivations of  $S$  are inner which implies that

$$S \otimes O(2; \underline{1}) \subset \bar{\mathcal{G}} = (S \otimes O(2; \underline{1})) \rtimes (\text{Id}_S \otimes \mathcal{D})$$

for some transitive Lie subalgebra of  $W(2; \underline{1})$ .

3.6. Suppose  $s = m = 2$ . Then  $\mathcal{G}_1 \subseteq \mathbb{k} \partial_1 \oplus \mathbb{k} \partial_2$  and

$$\mathcal{G}_0 \cong \bar{\mathcal{G}}_0 = (S \otimes 1) \oplus (\text{Id}_S \otimes \mathcal{D}_0) \subseteq (S \otimes 1) \oplus \left( \sum_{i,j=1}^2 \mathbb{k}(x_i \partial_j) \right)$$

where  $\mathcal{D}_0 = \mathcal{D} \cap W(2; \underline{1})[0]$ . If  $\dim \mathcal{G}_1 = 1$  we can argue as in the previous paragraph to conclude that  $M$  is contained in a proper parabolic subalgebra of  $\mathfrak{g}$ . So assume that  $\mathcal{G}_1 = \mathbb{k} \partial_1 \oplus \mathbb{k} \partial_2$ . Then  $\mathcal{D}_0 \subseteq \sum_{i,j=1}^2 \mathbb{k}(x_i \partial_j) \cong \mathfrak{gl}_2$  acts faithfully on  $\mathcal{G}_1$ . Let  $Q$  denote the centraliser of  $\text{nil}(M)$  in  $\mathfrak{g}$ . This is a restricted ideal of  $M$ . Since  $\mathcal{G}_2 = 0$ , the ideal  $\text{nil}(M)$  is abelian. Hence  $\text{nil}(M) \subseteq Q$ . Write  $\bar{Q}$  for the image of  $Q$  in  $\mathcal{G}_0 = M/\text{nil}(M)$ . Since  $S \otimes \mathbb{k} \subset \mathcal{G}_0$  commutes with  $\mathcal{G}_1$  and  $\mathcal{D}_0$  acts faithfully on  $\mathcal{G}_1$ , we now deduce that  $\bar{Q} = S \otimes \mathbb{k}$ .

Our next goal is to produce a nice subalgebra  $R$  of  $M$  satisfying the conditions of Lemma 2.2. If  $S \cong \mathfrak{sl}_2$  we choose a standard basis  $\{e_0, h_0, f_0\}$  of  $S$ . If  $S \cong W(1; \underline{1})$  we choose an  $\mathfrak{sl}_2$ -triple  $\{e_0, h_0, f_0\} \subset S \setminus \{0\}$  such that  $f_0 \in \mathbb{k}\partial$ ,  $h_0 = x\partial$  and  $e_0 \in \mathbb{k}(x^2\partial)$ . This choice ensures that  $\mathfrak{c}_S(f_0) = \mathbb{k}f_0$ . To ease notation we identify  $S$  with  $S \otimes \mathbb{k} \subset \bar{\mathcal{G}}_0$  and regard  $e_0, h_0, f_0$  as elements of  $\bar{\mathcal{G}}_0$ . Let  $\tilde{h}_0 \in Q$  be a preimage of  $h_0$  under the canonical homomorphism  $Q \twoheadrightarrow \bar{Q} = Q/M_{(1)}$ . By construction,  $h_0$  is a toral element of  $\mathcal{G}_0$ . Replacing  $\tilde{h}_0$  by its sufficiently large  $[p]$ -th power in  $Q$  we may assume that  $\tilde{h}_0$  is a semisimple element of  $\mathfrak{g}$ . Then the inverse image of  $f_0$  under the canonical homomorphism  $Q \twoheadrightarrow \bar{Q} = Q/M_{(1)}$  contains an eigenvector for  $\text{ad } \tilde{h}_0$ ; we call it  $\tilde{f}_0$ . Since  $f_0^{[p]} = 0$  in  $S$  it must be that  $\tilde{f}_0^{[p]} \in M_{(1)}$ . Therefore,  $\tilde{f}_0$  is a nilpotent element of  $\mathfrak{g}$ . We now set  $R := \mathbb{k}\tilde{h}_0 \oplus \mathbb{k}\tilde{f}_0 \oplus \text{nil}(M)$ . Clearly, this is Lie subalgebra of  $M$  and  $[\bar{R}, \bar{R}] \subseteq \mathbb{k}\tilde{f}_0 \oplus \text{nil}(M)$ . So  $[R, R]$  consists of nilpotent elements of  $\mathfrak{g}$  by Jacobson's formula for  $p$ -th powers. By Corollary 2.3, there exists a nonzero  $e \in Q$  such that  $[R, e] \subseteq \mathbb{k}e$  and  $[e, [e, \mathfrak{g}]] = \mathbb{k}e$ . Let  $\bar{e}$  denote the image of  $e$  in  $\bar{Q}$ .

Suppose  $\bar{e} \neq 0$ . As  $[f_0, \bar{e}] = 0$  we then have  $\bar{e} \in \mathbb{k}^\times f_0$ . So our assumption on  $\bar{e}$  entails  $\text{rk}(\text{ad } f_0)^2 \leq 1$ . However, the restriction of  $(\text{ad } f_0)^2$  to every subspace  $\bar{\mathcal{G}}_{-i} = S \otimes O(2; \underline{1})[i; 2]$  with  $1 \leq i \leq 2(p-1)$  is nonzero, forcing  $\text{rk}(\text{ad } e)^2 \geq \text{rk}(\text{ad } f_0)^2 \geq 2p-2$ . This contradiction shows that  $\bar{e} = 0$ . As a consequence,  $e \in M_{(1)}$ . Since  $\mathcal{G}_1 = \mathbb{k}\partial_1 \oplus \mathbb{k}\partial_2$ , we may assume after a suitable linear substitution of  $x_1, x_2$  that the image of  $e$  in  $\bar{\mathcal{G}}$  equals  $\partial_1$ . But since  $(\text{ad } e)^3 = 0$  this would entail  $(\text{ad } \partial_1)^3 = 0$  which is false because  $(\text{ad } \partial_1)^{p-1} \neq 0$  and  $p > 3$ . We thus conclude that the present case cannot occur.

3.7. Now suppose  $m = 2$  and  $s = 1$ , so that again  $S$  is either  $\mathfrak{sl}_2$  or  $W(1; \underline{1})$ . This case is quite similar to the previous one, but we need to choose a solvable subalgebra  $R$  more carefully. By (3.4),  $\mathcal{G}_1 \subseteq O(x_2)\partial_1$  and  $\mathcal{G}_0 \cong \bar{\mathcal{G}}_0 = (S \otimes O(x_2)) \oplus (\text{Id}_S \otimes \mathcal{D}_0)$  where

$$\mathcal{D}_0 = \mathcal{D} \cap W(2; \underline{1})[0; 1] \subseteq O(x_2)\partial_2 \oplus O(x_2)(x_1\partial_1).$$

Let  $\mathfrak{m}_2$  be the maximal ideal of the local ring  $O(x_2)$ . Thanks to the transitivity of  $\mathcal{D} \subseteq \mathcal{D}_0 \oplus \mathcal{G}_1$  and the inclusion  $\mathcal{G}_1 \subseteq O(x_2)\partial_1$  there exists an element  $d \in \bar{\mathcal{G}}_0$  such that  $\pi(d) = \phi(x_2)\partial_2 + \psi(x_2)(x_1\partial_1) \in \mathcal{D}_0$  has the property that  $\phi(x_2) \in O(x_2) \setminus \mathfrak{m}_2$ . Since all derivations of  $S$  are inner, our discussion in (3.4) yields  $d - \pi(d) \in S \otimes O(x_2)$ .

As before, we may assume that  $\dim \mathcal{G}_1 \geq 2$ . Indeed, otherwise  $\text{nil}(M) = M_{(1)}$  is 1-dimensional implying that  $M = \mathfrak{n}_{\mathfrak{g}}(M_{(1)})$  is contained in a parabolic subalgebra of  $\mathfrak{g}$ . We choose  $e_0, f_0, h_0 \in S$  as in (3.6) and define  $\bar{R} := \mathbb{k}d \oplus \mathbb{k}(h_0 \otimes 1) \oplus (f_0 \otimes O(x_2))$ . Then  $[R, R] \subseteq f_0 \otimes O(x_2)$  yielding  $[\bar{R}, \bar{R}]^{[p]} \subseteq f_0^{[p]} \otimes O(x_2) = 0$ . Let  $R$  be the preimage of  $\bar{R}$  under the canonical homomorphism  $M \twoheadrightarrow \mathcal{G}_0$ . By construction,  $R$  is a Lie subalgebra of  $M$  with  $[R, R]^{[p]} \subseteq M_{(1)}$ . Therefore, it satisfies the assumptions of Lemma 2.2.

Now we need to locate  $\bar{Q}$ , the image of  $Q = \mathfrak{c}_{\mathfrak{g}}(M_{(1)})$  in  $\mathcal{G}_0$ . It contains  $S \otimes O(x_2)$  because the latter commutes with  $\mathcal{G}_1 \subseteq O(x_2)\partial_1$ . Therefore,

$$\bar{Q} = (\text{Id}_S \otimes \text{Ann}_{\mathcal{D}_0}(\mathcal{G}_1)) \oplus (S \otimes O(x_2)).$$

Clearly,  $\text{Ann}_{\mathcal{D}_0}(\mathcal{G}_1)$  is an ideal of  $\mathcal{D}_0$ . Since  $\mathcal{G}_1 \subseteq O(x_2)\partial_1$  is  $d$ -stable and  $O(x_2)(x_1\partial_1)$  preserves  $\mathfrak{m}_2\partial_1 \subset \mathcal{G}_1$ , the subspace  $\mathcal{G}_1$  contains an element of the form  $a(x_2)\partial_1$  with  $a(x_2) \in O(x_2) \setminus \mathfrak{m}_2$ . Since

$$[f(x_2)(x_1\partial_1), a_2(x_2)\partial_1] = -f(x_2)a(x_2)\partial_1 \quad (\forall f(x_2) \in O(x_2))$$

and  $a(x_2)$  is invertible in  $O(x_2)$ , it must be that  $(O(x_2)(x_1\partial_1)) \cap \text{Ann}_{\mathcal{D}_0}(\mathcal{G}_1) = 0$ .

Suppose  $\text{Ann}_{\mathcal{D}_0}(\mathcal{G}_1) \neq 0$ . Then there exists  $u = b(x_2)\partial_2 + c(x_2)(x_1\partial_1) \in \text{Ann}_{\mathcal{D}_0}(\mathcal{G}_1)$  such that  $b(x_2), c(x_2) \in O(x_2)$  and  $b(x_2) \neq 0$ . Since  $O(x_2)(x_1\partial_1)$  is an abelian ideal of the Lie algebra  $O(x_2)\partial_2 \oplus O(x_2)(x_1\partial_1)$  and  $\text{Ann}_{\mathcal{D}_0}(\mathcal{G}_1)$  is  $(\text{ad } d)$ -stable, we may assume further that  $b(x_2) \notin \mathfrak{m}_2$ . As  $\dim \mathcal{G}_1 \geq 2$  and  $\mathfrak{m}_2\partial_1$  has codimension 1 in  $O(x_2)\partial_1$  it must be that  $(\mathfrak{m}_2\partial_1) \cap \mathcal{G}_1 \neq 0$ . Let  $v = g(x_2)\partial_1$  be a nonzero vector in  $(\mathfrak{m}_2\partial_1) \cap \mathcal{G}_1$ . Then there is  $k \in \{1, \dots, p-1\}$  such that  $g(x_2) \in \mathfrak{m}_2^k \setminus \mathfrak{m}_2^{k+1}$  and

$$[u, v] = [b(x_2)\partial_2 + c(x_2)(x_1\partial_1), g(x_2)\partial_1] = (b(x_2)g'(x_2) - c(x_2)g(x_2))\partial_1 \notin \mathfrak{m}_2^k\partial_1$$

(it is important here that  $b(x_2) \notin \mathfrak{m}_2$ ). In particular,  $[u, v] \neq 0$ . Since this contradicts our choice of  $u$  we now deduce that  $\text{Ann}_{\mathcal{D}_0}(\mathcal{G}_1) = 0$ . This yields  $\bar{Q} = S \otimes O(x_2)$ .

By Corollary 2.3, there exists a nilpotent element  $e \in Q$  such that  $[R, e] \subseteq \mathbb{k}e$  and  $[e, [e, \mathfrak{g}]] = \mathbb{k}e$ . Let  $\bar{e}$  be the image of  $e$  in  $\bar{Q} = S \otimes O(x_2)$ . Since  $[f_0 \otimes 1, \bar{e}] = 0$  it must be that  $\bar{e} = f_0 \otimes q(x_2)$  for some  $q(x_2) \in O(x_2)$ . Suppose  $\bar{e} \neq 0$ . Note that  $O(x_2)(x_1\partial_1)$  acts trivially on  $S \otimes O(x_2)$ . Since

$$[d, f_0 \otimes q(x_2)] \in \mathbb{k}(f_0 \otimes q(x_2)) \cap (f_0 \otimes \phi(x_2)q'(x_2) + S \otimes O(x_2)q(x_2))$$

and  $\phi(x_2) \notin \mathfrak{m}_2$ , it must be that  $q(x_2) \notin \mathfrak{m}_2$ . From this it is immediate that the restriction of  $(\text{ad } \bar{e})^2$  to every subspace  $\bar{\mathcal{G}}_{-i} = S \otimes O(2; \underline{1})[i; 1]$  with  $1 \leq i \leq p-1$  is nonzero. But then  $\text{rk}(\text{ad } e)^2 \geq \text{rk}(\text{ad } \bar{e})^2 \geq p-1$ . This contradiction shows that  $e \in M_{(1)}$ . Identifying  $M_{(1)}$  with  $\mathcal{G}_1 \subseteq O(x_2)\partial_1$  we write  $e = w(x_2)\partial_1$  for some nonzero  $w(x_2) \in O(x_2)$ . Since  $\phi(x_2) \notin \mathfrak{m}_2$  and

$$[d, w(x_2)\partial_1] = (\phi(x_2)w'(x_2) - \psi(x_2)w(x_2))\partial_1 \in \mathbb{k}(w(x_2)\partial_1)$$

it is straightforward to see that  $w(x_2) \notin \mathfrak{m}_2$ . But then  $(w(x_2)\partial_1)^{p-1}$  is a nonzero linear operator on  $O(2; \underline{1})$  implying that  $(\text{ad } e)^3 \neq 0$ . Since this is false, we finally conclude that the degenerate case of Weisfeiler's theorem cannot occur for  $\bar{\mathcal{G}}$ .

3.8. From now on we may assume that the non-degenerate case of Weisfeiler's theorem holds for  $\bar{\mathcal{G}}$ , that is  $S = \bigoplus_{i \in \mathbb{Z}} S_i$  is  $\mathbb{Z}$ -graded,  $A_i(\bar{\mathcal{G}}) = S_i \otimes O(m; \underline{n})$  for all  $i \in \mathbb{Z}$ , and

$$S_0 \otimes O(m; \underline{n}) \subseteq \bar{\mathcal{G}}_0 \subseteq (\text{Der}_0(S) \otimes O(m; \underline{n})) \rtimes (\text{Id}_S \otimes \mathcal{D})$$

where  $\mathcal{D} = \pi(\bar{\mathcal{G}}_0)$  is a transitive subalgebra of  $W(m; \underline{n})$ . We shall often identify  $\mathcal{G}_{\geq 0}$  with  $\bar{\mathcal{G}}_{\geq 0}$ .

First suppose that  $S = \text{Lie}(\mathcal{H})$  for some simple algebraic  $\mathbb{k}$ -group  $\mathcal{H}$  (this excludes the case where  $S \cong \mathfrak{psl}_{kp}$  for some  $k \geq 1$ ). Then  $\text{Der}(S) = \text{ad } S$ ; see [BGP09, Lemma 2.7], for example. It follows that  $\bar{\mathcal{G}}_i = S_i \otimes O(m; \underline{n})$  for all  $i \neq 0$ . Our grading of  $S$  is induced by the action of a 1-dimensional torus  $T_0$  of  $(\text{Aut } S)^\circ = \text{Ad } \mathcal{H}$ . Differentiating this action we find a toral element  $t_0 \in \text{Lie}(T_0)$  such that  $[t_0, x] = \bar{i}x$  for all  $x \in S_i$  where  $i \in \mathbb{Z}$ . Here and in what follows we write  $\bar{i}$  for the image of  $i$  in  $\mathbb{F}_p \subset \mathbb{k}$ . The element  $t_0 \in \text{Der}(S)$  is often referred to as the *degree derivation* of the graded Lie algebra  $S$ . In the present case the derivation  $t_0$  is inner.

It is well known that  $S_{\geq 0} = \bigoplus_{i \geq 0} S_i$  is a parabolic subalgebra of  $S$  and  $S_0 = S^{T_0}$  is a Levi subalgebra of  $S_{\geq 0}$ . Obviously,  $t_0 \in S_0$  and we may assume without loss of generality that  $S_{\geq 0}$  is a standard parabolic subalgebra of  $S$ . Since  $M_{(0)} = M$  is a restricted subalgebra of  $\mathfrak{g}$ , there exists a toral element  $\tilde{t}_0 \in M_{(0)}$  which maps onto  $t_0$  under the canonical homomorphism  $M_{(0)} \twoheadrightarrow \mathcal{G}_0$ . If  $\mathfrak{c}_{\mathcal{G}}(t_0) \subseteq \mathcal{G}_{\geq 0}$  then  $\mathfrak{g}_{\tilde{t}_0} \subseteq M_{(0)}$ . Since  $\mathfrak{g}_{\tilde{t}_0}$  is a Levi subalgebra of  $\mathfrak{g}$ , it contains a maximal torus of  $\mathfrak{g}$ . This, however, contradicts our assumption that  $M$  is not a regular subalgebra of  $\mathfrak{g}$ . So it must be that  $\mathfrak{c}_{\mathcal{G}}(t_0) \cap \mathcal{G}_i \neq 0$  for some  $i < 0$ .

Since  $\text{Der}(S) = \text{ad } S$ , our discussion at the end of (3.4) shows that the  $S_0$ -module  $S_{-1}$  is semisimple and isogenic. In particular,  $\mathfrak{z}(S_0)$  acts on  $S_{-1}$  by scalar endomorphisms forcing  $\mathfrak{z}(S_0) = \mathbb{k}t_0$ . From

this it follows that the parabolic subalgebra  $S_{\geq 0}$  is maximal in  $S$  and our grading of  $S$  is standard. This means that there exist maximal torus  $T$  of  $\mathcal{H}$  containing  $T_0$ , a basis of simple roots  $\{\alpha_1, \dots, \alpha_\ell\}$  in the root system  $\Phi(\mathcal{H}, T)$ , and a positive integer  $d \leq \ell$  such that  $S_k$  with  $k \neq 0$  is spanned by all root vectors  $e_\gamma$  of  $S$  with respect to  $T$  such that  $\gamma = \sum_{i=1}^\ell m_i \alpha_i$  and  $m_d = k$ ; see [BGP09, 2.4] for detail.

Note that  $p > 3$  is good for  $S$  provided that  $S$  is not of type  $E_8$  and if  $S$  is of type  $E_8$  then so is  $\mathfrak{g}$ . So in any event  $p$  is good for  $S$ . Since our grading of  $S$  is standard we now obtain that  $\bar{\mathcal{G}}_+ = \bigoplus_{i>0} (S_i \otimes O(m; \underline{n}))$  is generated by  $\bar{\mathcal{G}}_1 = S_1 \otimes O(m; \underline{n})$ . Furthermore,  $\dim S_i = \dim S_{-i}$  for all  $i \in \mathbb{Z}$ , and  $S_i = 0$  for all  $i \geq p$ . This yields  $\mathfrak{c}_{\bar{\mathcal{G}}}(t_0) \subseteq \bar{\mathcal{G}}_{\geq 0}$ . Since  $\mathfrak{c}_{\bar{\mathcal{G}}}(t_0) \cap \mathcal{G}_i \neq 0$  for some  $i < 0$  it follows that  $N(\bar{\mathcal{G}}) \neq 0$ .

Let  $l$  be the smallest positive integer for which  $N_{-l}(\mathcal{G}) \neq 0$ . Then  $2 \leq l \leq p$  and  $[N_{-l}(\mathcal{G}), \mathcal{G}_1] = 0$ . Since  $\mathfrak{g}$  is generated by  $M_{(-1)}$ , the toral element  $\tilde{t}_0$  preserves each  $M_{(i)}$  with  $i \in \mathbb{Z}$  and acts on  $M_{(i)}/M_{(i+1)}$  as the scalar operator  $\bar{i} \cdot \text{Id}$ . Since  $\text{ad } \tilde{t}_0$  is diagonalisable, for every  $i \in \mathbb{Z}$  there exists a subspace  $V_i \subset \mathfrak{g}$  such that  $[\tilde{t}_0, v] = \bar{i}v$  for all  $v \in V_i$  and  $M_{(i)} = V_i \oplus M_{(i+1)}$ . Given  $a \in \mathbb{k}$  and an  $(\text{ad } \tilde{t}_0)$ -invariant subspace  $W$  of  $\mathfrak{g}$  we write  $W(a)$  for the set of all  $w \in W$  with  $[\tilde{t}_0, w] = aw$ .

Let  $\bar{u} \in N_{-l}(\mathcal{G}) \setminus \{0\}$  and pick  $u \in V_{-l}$  which maps onto  $\bar{u}$  under the canonical homomorphism  $M_{(-l)} \twoheadrightarrow \mathcal{G}_{-l} = M_{(-l)}/M_{(-l+1)}$ . Since  $[\bar{u}, \mathcal{G}_1] = 0$  it must be that  $[u, M_{(1)}] \subseteq M_{(-l+2)}$ . But then

$$[u, V_1] \subseteq M_{(-l+2)}(\overline{-l+1}) = (V_{-l+2} \oplus \dots \oplus V_{-1} \oplus V_0 \oplus M_{(1)}) \cap \mathfrak{g}(\overline{-l+1}) = M_{(1)}(\overline{-l+1}).$$

(It is of utmost importance here that the set of residues of  $0, -1, \dots, -l+2$  in  $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$  does not contain the residue of  $-l+1$ ). This shows that  $[u, V_1] \subseteq M_{(1)}$ . Since  $V_1$  maps onto  $\mathcal{G}_1$  under the canonical homomorphism  $M_{(1)} \twoheadrightarrow \mathcal{G}_1$  and  $\mathcal{G}_1$  generates the Lie algebra  $\mathcal{G}_+$  by our remarks earlier in the proof, the nilpotent Lie algebra  $M_{(1)}$  must be generated by  $V_1$ . This gives  $[u, M_{(1)}] \subseteq M_{(1)}$  forcing  $u \in \mathfrak{n}_{\mathfrak{g}}(M_{(1)})$ . But then  $u \in M$  by the maximality of  $M$ . This contradiction shows that the case where  $S = \text{Lie}(\mathcal{H})$  is impossible.

3.9. In the next four subsections we assume that  $S = \mathfrak{psl}_{kp}$  for some  $k \in \mathbb{Z}_{>0}$ . Recall that for a finite dimensional Lie algebra  $L$  over  $\mathbb{k}$  the *absolute toral rank*  $\text{TR}(L)$  denotes the maximal dimension of toral subalgebras in the restricted Lie algebras  $\tilde{L}/\mathfrak{z}(\tilde{L})$  where  $\tilde{L}$  runs over the set of all finite dimensional  $p$ -envelopes of  $L$ . Since  $p > 2$ , the restricted Lie algebra  $\mathfrak{psl}_{kp}$  contains a self-centralising torus of dimension  $kp - 2$ . Similarly,  $\mathfrak{g}$  contains a self-centralising torus whose dimension equals  $\text{rk}(\mathfrak{g})$ . So, combining [Pre87] and [Pre90] with [Str04, Theorem 1.2.9] one obtains that  $\text{TR}(\mathfrak{g}) = \text{rk}(\mathfrak{g})$  and  $\text{TR}(\mathfrak{psl}_{kp}) = kp - 2$ . On the other hand, Skryabin's theorem says that that  $\text{TR}(\mathfrak{g}) \geq \text{TR}(\mathcal{G})$ ; see [Skr98, Theorem 5.1]. By [Str04, Theorems 1.2.8(3) and 1.2.7(1)], this gives

$$kp - 2 = \text{TR}(\mathfrak{psl}_{kp}) = \text{TR}(S) \leq \text{TR}(\bar{\mathcal{G}}) \leq \text{TR}(\mathcal{G}) \leq \text{TR}(\mathfrak{g}) = \text{rk}(\mathfrak{g}) \leq 8.$$

Since  $p > 5$  when  $G$  is of type  $E_8$ , this yields  $k = 1$ . Since  $\dim \mathfrak{g} > (p^2 - 2)p^{|\underline{n}|}$  where  $|\underline{n}| = n_1 + \dots + n_m$ , it must be that  $m = 0$  unless  $\mathfrak{g}$  is of type  $E_7$ ,  $p = 5$ , and  $O(m; \underline{n}) = O(1, \underline{1})$ .

The grading of  $S$  is induced by the action of a 1-dimensional torus  $\lambda(\mathbb{k}^\times)$  of  $(\text{Aut } S)^\circ \cong \text{PGL}_p$  on  $S$  and hence  $S_{\geq 0}$  is a parabolic subalgebra of  $S$ . However, it is no longer true in the present case that the degree derivation  $t_0 \in \text{Lie}(\lambda(\mathbb{k}^\times))$  lies in  $\text{ad } S_0$ . Essentially this is due to the fact that  $\text{Der}(\mathfrak{psl}_p) \cong \mathfrak{pgl}_p$ ; see [BGP09, Lemma 2.7], for example. Still  $\lambda(\mathbb{k}^\times)$  is contained in a maximal torus  $T$  of  $\text{PGL}_p$  and there is a basis of simple roots  $\Delta$  in the root system of  $\text{Aut } S$  with respect to  $T$  and a collection of non-negative integers  $\{r_\alpha \mid \alpha \in \Delta\}$  such that  $\alpha(\lambda(t)) = t^{r_\alpha}$  for all  $\alpha \in \Delta$  and all  $t \in \mathbb{k}^\times$ .

If  $m = 0$  then [Str04, Theorem 3.5.7.(6)] shows that the  $S_0$ -module  $S_{-1}$  is semisimple and isogenic. If  $m > 0$  then  $S_0 = \mathfrak{c}_S(\lambda)$  is a proper standard Levi subalgebra of  $S \cong \mathfrak{psl}_5$ . Let  $\Sigma_0$  be the root system of  $S_0$  with respect to  $T$ . Since  $\text{Der}(S)$  has type  $A_4$ , the Lie algebra  $S_0$  is either toral or  $\Sigma_0$  has type  $A_1$ ,  $A_1^2$ ,  $A_2$ ,  $A_2A_1$  or  $A_3$ . In any event, this implies that  $S_{-1}$  is a semisimple  $S_0$ -module and all its irreducible submodules have dimension 1, 2, 3, 4 or 6. Moreover, in the last two cases  $S_{-1}$  is an irreducible  $S_0$ -module. If  $S_0$  is toral then  $S_{-1}$  coincides with the span of the roots vectors  $e_{-\alpha} \in S$  such that  $\alpha \in \Delta$  and  $r_\alpha = 1$ . In particular,  $\dim S_{-1} \leq 4$ . Taking all this into account it is not hard observe that  $S_{-1}$  always contains an irreducible  $S_0$ -submodule  $U_0$  such that  $\dim S_{-1} < p \dim U_0$ . Applying [Str04, Theorem, 3.5.7(6)] we now deduce that when  $m > 0$  the  $S_0$ -module  $S_{-1}$  is semisimple and isogenic, too. This implies that in all cases of interest the parabolic subalgebra  $S_{\geq 0}$  is maximal in  $S$ . From this it is immediate that  $S = S_{-1} \oplus S_0 \oplus S_1$ .

3.10. The above discussion in (3.9) also shows that  $\text{Der}(S) = \mathbb{k}t_0 \oplus \text{ad } S$  and

$$S_0 \otimes O(m; \underline{n}) \subseteq \mathcal{G}_0 \subseteq (\text{Der}_0(S) \otimes O(m; \underline{n})) \rtimes (\text{Id}_S \otimes \mathcal{D})$$

for some transitive subalgebra  $\mathcal{D}$  of  $W(m; \underline{n})$ . Since all Cartan subalgebras of  $\mathfrak{sl}_p$  are toral the subalgebra  $S_0$  of  $S = \mathfrak{sl}_p/\mathbb{k}1$  contains a toral Cartan subalgebra of  $S$ , say  $\mathfrak{t}$ . Clearly,  $\mathfrak{t}$  is a toral subalgebra of  $\mathcal{G}_0 = \bar{\mathcal{G}}_0$ . Since the canonical homomorphism  $\eta: M_{(0)} \rightarrow \mathcal{G}_0$  is restricted, there exists a toral subalgebra  $\tilde{\mathfrak{t}}$  in  $M_{(0)}$  such that  $\eta(\tilde{\mathfrak{t}}) \cong \mathfrak{t}$ ; see [Str04, Proposition 1.2.2], for example.

Since  $\mathfrak{t} \subset S_0$  is self-centralising we have that  $\mathfrak{c}_S(\mathfrak{t}) \cap S_{-1} = 0$ . Consequently,  $\mathfrak{c}_{\bar{\mathcal{G}}}(\mathfrak{t}) \cap \bar{\mathcal{G}}_{-1} = 0$ . If  $\mathcal{G}_{-2} = 0$  then  $\mathfrak{c}_{\mathcal{G}}(\mathfrak{t}) \subset \mathcal{G}_{\geq 0}$  forcing  $\mathfrak{c}_{\mathfrak{g}}(\tilde{\mathfrak{t}}) \subset M_{(0)} = M$ . As this contradicts our assumption that  $M$  is non-regular we now deduce that  $\mathcal{G}_{-2} = N_{-2}(\mathcal{G}) \neq 0$ . Since  $N(\mathcal{G}) \subseteq \bigoplus_{i \leq -2} N_i(\mathcal{G})$  we have that  $\bar{\mathcal{G}}_1 \cdot \bar{N}_{-2}(\mathcal{G}) = 0$ . The graded factor-space  $\bar{N}(\mathcal{G}) := N(\mathcal{G})/N(\mathcal{G})^2$  carries a natural  $\bar{\mathcal{G}}$ -module structure and  $\bar{N}_{-i}(\mathcal{G}) \cong N_{-i}(\mathcal{G})$  for  $i = 2, 3$  as vector spaces. Therefore,  $\bar{N}_{-2}(\mathcal{G}) \neq 0$ . If  $\bar{\mathcal{G}}_{-1} \cdot \bar{N}_{-2}(\mathcal{G}) = 0$  then  $\mathcal{G}_{-3} = [\mathcal{G}_{-1}, \mathcal{G}_{-2}] = N_{-3}(\mathcal{G}) = 0$  because  $N_{-3}(\mathcal{G}) \cong \bar{N}_{-3}(\mathcal{G})$ .

Suppose  $\bar{\mathcal{G}} \cdot \bar{N}(\mathcal{G}) = 0$ . Then the above shows that  $N(\mathcal{G}) = \mathcal{G}_{-2}$  and  $[\mathcal{G}_{\pm 1}, N(\mathcal{G})] = 0$ . It is well known that the torus  $\tilde{\mathfrak{t}} \subset M_{(0)}$  is contained in a maximal torus of  $\mathfrak{g}$  and all such tori are  $(\text{Ad } G)$ -conjugate by [Bor91, 11.8] or [Hum67, Theorem 13.3]. Since  $\mathfrak{t}$  is a maximal torus of  $S_0$  we have that  $\dim \tilde{\mathfrak{t}} = p - 2$ . If  $p \geq 7$  then  $p - 2 > \frac{1}{2} \text{rk}(\mathfrak{g})$ . In this case, the restriction of the (non-degenerate) Killing form of  $\mathfrak{g}$  to  $\tilde{\mathfrak{t}}$  is nonzero. Since  $\tilde{\mathfrak{t}} \subset M_{(0)}$  and  $\mathfrak{t} \subset S_0 \subseteq [\mathcal{G}_{-1}, \mathcal{G}_1]$ , one of the composition factors of the  $S$ -module  $\bar{\mathcal{G}} = \mathcal{G}/\mathcal{G}_{-2}$  provides  $S$  with a nonzero trace form. But since  $S \cong \mathfrak{psl}_p$  no such form can exist by [Bl62] or [Gar09]. This shows that if  $\bar{\mathcal{G}} \cdot \bar{N}(\mathcal{G}) = 0$  then  $p = 5$  and  $\mathfrak{g}$  is not of type  $E_8$ . Arguing similarly one also observes that  $G$  cannot be of type  $G_2$  or  $F_4$  (even when  $p = 5$ ).

3.11. Suppose  $m = 0$ . Then  $S_0 \subseteq \mathcal{G}_0 \subseteq \text{Der}_0(S) = S_0 \oplus \mathbb{k}t_0$ . If  $t_0 \in \mathcal{G}_0$  then arguing as in (3.8) we could find a toral element  $\tilde{t}_0 \in M_{(0)}$  with  $\mathfrak{g}_{\tilde{t}_0} \subseteq M$ . Since this contradicts our assumption that  $M$  is non-regular, the equality  $\bar{\mathcal{G}} = S$  must hold. In particular, this means that  $\mathcal{G}_1 = S_1$  and  $\mathcal{G}_0 = S_0$  acts faithfully on  $\mathcal{G}_1$ . From this it is immediate that the subalgebra  $\mathfrak{c}_{\mathfrak{g}}(\text{nil}(M))$  of  $M = M_{(0)}$  coincides with  $M_{(1)}$ . Then there exists a nonzero  $e \in M_{(1)}$  such that  $[e, [e, \mathfrak{g}]] = \mathbb{k}e$ ; see Corollary 2.3. Its image  $\bar{e} = \text{gr}_1(e)$  in  $\mathcal{G}_1 = S_1$  has the property that  $[\bar{e}, [\bar{e}, \mathcal{G}_i]] = 0$  for all  $i \leq -2$ .

Since  $\bar{\mathcal{G}} = S$  we have that  $\bar{\mathcal{G}} = \bar{\mathcal{G}}^{(\infty)}$ . If all composition factors of the  $\bar{\mathcal{G}}$ -module  $\bar{N}(\mathcal{G})$  are trivial then  $\bar{\mathcal{G}} \cdot \bar{N}(\mathcal{G}) = \bar{\mathcal{G}}^{(\infty)} \cdot \bar{N}(\mathcal{G}) = 0$ . Then  $\bar{\mathcal{G}}_{-1} \cdot \bar{N}_{-2}(\mathcal{G}) = 0$  and our discussion in (3.10) shows that  $p = 5$ . Since  $\mathcal{G}_{-2} = [\mathcal{G}_{-1}, \mathcal{G}_{-1}]$  is a homomorphic image of  $\wedge^2 \mathcal{G}_{-1}$  and  $\dim \mathcal{G}_{-1} = k(5 - k)$  for some  $1 \leq k \leq 4$  we have that  $\dim \mathcal{G}_{-2} \leq 15$ . But then  $\dim \mathfrak{g} = \dim S + \dim \mathcal{G}_{-2} \leq 23 + 15 = 48$ , a contradiction.



Let  $V$  be a non-trivial composition factor of the  $S$ -module  $\bar{N}(\mathcal{G})$  and denote by  $\rho: S \rightarrow \mathfrak{gl}(V)$  the corresponding representation of  $S = \mathfrak{psl}_p$ . We shall regard  $\rho$  as a representation of  $\mathfrak{sl}_p = \text{Lie}(\text{SL}_p)$  by inflation. Our earlier remarks then show that  $\rho(a)^2 = 0$  for some non-central element  $a \in \mathfrak{sl}_p$  (one can take  $a = \bar{e}$  since  $(\text{ad}_{\mathcal{G}} \bar{e})^2 = 0$ ). It should be stressed at this point that the centre  $\mathbb{k}I_p$  of  $\mathfrak{sl}_p$  acts trivially on  $V$ . By Lemma 2.5,  $\rho$  is a restricted representation of  $S$ . So, thanks to Curtis' theorem it can be obtained by differentiating an infinitesimally irreducible rational representation of  $\text{SL}_p$ . Since the  $(\text{Ad } \text{SL}_p)$ -stable set  $\{X \in \mathfrak{sl}_p \mid \rho(X)^2 = 0\}$  is non-central, Zariski closed and conical, it contains the orbit  $\mathcal{O}_{\min}$  of  $\mathfrak{sl}_p$ . This shows that  $\rho(e_\alpha)^2 = 0$  for any root vector  $e_\alpha \in \mathfrak{sl}_p$ . Arguing as in [PSu83, p. 72] one then observes that  $\rho$  is a fundamental representation of  $\mathfrak{sl}_p$ , i.e.  $V \cong L(\varpi_k)$  for some  $1 \leq k \leq p-1$ . But then the central element  $I_p \in \mathfrak{sl}_p$  acts on  $V$  as  $k \cdot \text{Id}_V \neq 0$ , a contradiction. As a result,  $m \geq 1$ .

3.12. Our discussion in (3.11) implies that  $p = 5$ ,  $\mathfrak{g}$  is of type  $E_7$  and  $O(m; \underline{n}) = O(1; \underline{1})$ . Then

$$\dim \bar{N}(\mathcal{G}) \leq \dim N(\mathcal{G}) = \dim \mathfrak{g} - \dim \bar{\mathcal{G}} < \dim \mathfrak{g} - \dim S \otimes O(1; \underline{1}) = 133 - 115 = 18.$$

We claim that  $\mathcal{G}_{-3} = 0$ . Indeed, if this is not the case then  $\bar{\mathcal{G}}_{-1} \cdot \bar{N}_{-2}(\mathcal{G}) \neq 0$ . Since  $A(\bar{\mathcal{G}}) = S \otimes O(1; \underline{1})$  is a perfect Lie algebra, this implies that one of the composition factors of the  $A(\bar{\mathcal{G}})$ -module  $\bar{N}(\mathcal{G})$  is non-trivial. This, in turn, shows that there is a composition factor  $W$  of the  $\bar{\mathcal{G}}$ -module  $\bar{N}(\mathcal{G})$  such that  $A(\bar{\mathcal{G}}) \cdot W = W$ . In this situation Block's theorem on derivation simple modules says that there exists a faithful  $S$ -module  $W_0$  such that  $W = W_0 \otimes O(1; \underline{1})$  as vector spaces; see [Str04, Corollary 3.3.7]. As  $\dim W \leq \dim \bar{N}(\mathcal{G}) < 18$  and  $p = 5$  we get  $\dim W_0 \leq 3$ . But then  $23 = \dim S \leq \dim \mathfrak{gl}(W_0) = 9$ . By contradiction the claim follows.

Let  $M'$  denote the inverse image of  $S_0 \otimes O(1; \underline{1})$  under the canonical homomorphism  $M_{(0)} \twoheadrightarrow \mathcal{G}_0$ . Since  $\dim S_{-1} = k(5 - k)$  for some  $1 \leq k \leq 4$ , we have that  $\dim S_{\geq 0} \geq 23 - 6 = 17$ . As a consequence,  $\dim M' = \dim \mathcal{G}_{\geq 0} > \dim (S_{\geq 0} \otimes O(1; \underline{1})) \geq 85 > (\dim \mathfrak{g})/2$ . It follows that the restriction of the Killing form  $\kappa$  of  $\mathfrak{g}$  to  $M$  is nonzero. Since  $\text{ad } M_{(1)}$  acts nilpotently on  $\mathfrak{g}$ , it lies in the radical of the trace form  $\kappa|_{M'}$ . Let  $\kappa_L$  denote the Killing form of a finite dimensional Lie algebra  $L$ . Since  $M = M_{(0)}$  preserves each component  $M_{(i)}$  of our filtration, we have that

$$(9) \quad \kappa(x, y) = \text{tr}((\text{ad } x) \circ (\text{ad } y)) = \kappa_{\mathcal{G}}(\text{gr}_0(x), \text{gr}_0(y)) \quad (\forall x, y \in M).$$

In view of the above this shows that the restriction of  $\kappa_{\mathcal{G}}$  to  $S_0 \otimes O(1; \underline{1})$  is nonzero. As  $[\mathcal{G}_{\pm 1}, \mathcal{G}_{-2}] = 0$  by our earlier remarks and  $S_0 \otimes O(1; \underline{1}) = [\mathcal{G}_{-1}, \mathcal{G}_1]$ , we have that  $\kappa_{\mathcal{G}}(x, y) = \kappa_{\bar{\mathcal{G}}}(x, y)$  for all  $x, y \in S \otimes O(1; \underline{1})$  (as before we identify  $\bar{\mathcal{G}}_0$  with  $\mathcal{G}_0$ ). This implies that the restriction of  $\kappa_{\bar{\mathcal{G}}}$  to  $A(\bar{\mathcal{G}}) = S \otimes O(1; \underline{1})$  of  $\bar{\mathcal{G}}$  is nonzero. Since  $A(\bar{\mathcal{G}})$  is the unique minimal ideal of  $\bar{\mathcal{G}}$  and the form  $\kappa_{\bar{\mathcal{G}}}$  is  $\bar{\mathcal{G}}$ -invariant, the restriction of  $\kappa_{\bar{\mathcal{G}}}$  to  $A(\bar{\mathcal{G}})$  must be non-degenerate. On the other hand, it is straightforward to see that for any  $s \in S$  and any  $u \in A(\bar{\mathcal{G}})$  the linear operator  $\text{ad}_{A(\bar{\mathcal{G}})}(s \otimes x_1^{p-1}) \circ \text{ad}_{A(\bar{\mathcal{G}})} u$  is nilpotent and hence has zero trace. This contradiction shows that the case where  $S = \mathfrak{psl}_{kp}$  is impossible.

3.13. Suppose  $S$  is a Lie algebra of Cartan type  $W$ . Since  $\dim W(r; \underline{d}) = rp^{|\underline{d}|}$  and  $\dim S \leq \dim \mathfrak{g}$ , the Lie algebra  $S$  is in the following list:

$$W(1; \underline{1}), W(1; \underline{2}), W(1; \underline{3}), W(2; \underline{1}).$$

If  $S = W(1; \underline{d})$  for  $1 \leq |\underline{d}| \leq 3$  then we may assume that its grading is either natural or reverse thereof; see [PS01, Theorem 4.7]. In the second case,  $S_{-2} \neq 0$  and  $\dim S_{-1} = 1$ . But then

$$0 \neq S_{-2} \otimes O(m; \underline{n}) = \bar{\mathcal{G}}_{-2} = [\bar{\mathcal{G}}_{-1}, \bar{\mathcal{G}}_{-1}] = [S_{-1}, S_{-1}] \otimes O(m; \underline{n}) = 0,$$

a contradiction. Thus,  $S = \bigoplus_{i \geq -1} S_i$  and all graded components of  $S$  are 1-dimensional. Moreover, there exists  $t_0 \in S_0$  such that  $[t_0, x] = \bar{i}x$  for all  $x \in S_i$ . It follows that  $N(\mathcal{G}) = \bigoplus_{i \leq -2} \mathcal{G}_i$ , forcing  $[\mathcal{G}_{-2}, \mathcal{G}_1] = 0$ . Equivalently,  $[M_{(-2)}, M_{(i)}] \subseteq M_{(i-1)}$  for all  $i \geq 0$ .

The subspace  $\mathbb{k}(t_0 \otimes 1)$  is a 1-dimensional torus contained in  $\mathfrak{z}(\mathcal{G}_0)$ . As both  $\mathcal{G}_0$  and  $M_{(0)}$  are restricted, there exists a toral element  $\tilde{t}_0 \in M_{(0)}$  which maps onto  $t_0$  under the canonical epimorphism  $M_{(0)} \twoheadrightarrow \mathcal{G}_0$ . If  $N(\mathcal{G}) = 0$  then  $\mathfrak{g}_{\tilde{t}_0} \subset M_{(0)}$  contrary to our assumption that  $M = M_{(0)}$  is non-regular. This shows that  $\mathcal{G}_{-2} \neq 0$ . As in (3.8) we choose a subspace  $V_i$  in  $M_{(i)}$  such that  $M_{(i)} = V_i \oplus M_{(i+1)}$  and  $[\tilde{t}_0, x] = \bar{i}x$  for all  $x \in V_i$ . By our preceding remarks,  $V_{-2} \neq 0$  and  $[V_{-2}, V_i] \subseteq M_{(i-1)} \cap M_{(i-2)} \subset M_{(1)}$  for  $i = 1, 2$ . Since  $[V_{-2}, M_{(3)}] \subseteq M_{(1)}$  and  $M_{(1)} = V_1 \oplus V_2 \oplus M_{(3)}$  we thus deduce that  $0 \neq V_{-2} \subset \mathfrak{n}_{\mathfrak{g}}(M_{(1)}) = M_{(0)}$ . This contradiction shows that  $S \neq W(1; \underline{d})$ .

3.14. Suppose  $S = W(2; \underline{1})$ , a Lie algebra of dimension  $2p^2$ . Then  $m = 0$  and  $\text{Der}(S) = \text{ad } S$  by [Str04, Theorem 7.1.2(1)]. Hence  $\bar{\mathcal{G}} = S$ . In particular, this implies that  $S_{-1}$  is an irreducible and faithful  $S_0$ -module. By [PS01, Theorem 4.7], any grading of  $W(2; \underline{1})$  is given by assigning to some generators  $u_1, u_2$  of the maximal ideal of  $O(2; \underline{1})$  certain degrees  $a_1, a_2 \in \mathbb{Z}$ . The element  $u_1^{m_1} u_2^{m_2} \frac{\partial}{\partial u_k} \in W(2; \underline{1})$  with  $k = 1, 2$  then acquires the degree  $a_1 m_1 + a_2 m_2 - a_k$ . The grading of  $S$  obtained this way is said to have type  $(a_1, a_2)$  with respect to  $u_1, u_2$ . To ease notation we assume that  $u_i = x_i$  for  $i = 1, 2$ . We write  $S[i]$  for the  $i$ -th component of the standard grading of the Cartan type Lie algebra  $S$  which has type  $(1, 1)$ .

Suppose our grading of  $S$  has type  $(a_1, a_2)$ . No generality will be lost by assuming that  $|a_1| \geq |a_2|$ . Then  $S_0$  is spanned by some  $x_1^{i_1} x_2^{i_2} \partial_k$  with  $k = 1, 2$ . We now follow very closely the argument in [PS01, Lemma 4.15]. If  $\{\partial_1, \partial_2, x_1 \partial_2\} \cap S_0 = \emptyset$  then

$$\mathbb{k}x_1 \partial_1 \oplus \mathbb{k}x_2 \partial_2 \subseteq S_0 \subseteq \mathbb{k}x_1 \partial_1 \oplus \mathbb{k}x_2 \partial_2 \oplus \mathbb{k}x_2 \partial_1 \oplus \sum_{i \geq 1} S[i].$$

In this case  $S_0$  is solvable and  $[S_0, S_0] \subseteq \mathbb{k}x_2 \partial_1 \oplus \sum_{i \geq 1} S[i]$  acts nilpotently on  $S$ . Since  $S_{-1}$  is an irreducible and faithful  $S_0$ -module, this is impossible. Therefore,  $\{\partial_1, \partial_2, x_1 \partial_2\} \cap S_0 \neq \emptyset$ , implying that either  $a_2 = 0$  or  $a_1 = a_2$ . Since  $S_{-1} \neq 0$  we now deduce that either the grading of  $S$  or its reverse has type  $(1, 1)$  or  $(1, 0)$ . As before, we set  $Q := \mathfrak{c}_{\mathfrak{g}}(M_{(1)})$ , pick  $e \in Q \cap \mathcal{O}_{\min}$ , and denote by  $\bar{Q}$  the image of  $Q$  in  $\mathcal{G}_0 = M_{(0)}/M_{(1)}$ . Write  $\bar{e}$  for the image of  $e$  in  $\bar{Q}$ .

If the grading of  $S$  has type  $(-1, -1)$  then  $\mathcal{G}_1 = S_1 = \mathbb{k}\partial_1 \oplus \mathbb{k}\partial_2$  and  $S_i = 0$  for  $i \geq 2$ . Furthermore,  $\mathcal{G}_0 = S_0 = \sum_{i,j=1}^2 \mathbb{k}(x_i \partial_j)$  acts on  $S_1$  faithfully. From this it is immediate that  $\bar{Q} = 0$ . But then  $\bar{e} = \text{gr}_1(e)$  is a nonzero linear combination of  $\partial_1$  and  $\partial_2$ . Since  $(\text{ad } \bar{e})^3 = 0$  and  $p > 3$ , this is impossible. So this case cannot occur. If the grading of  $S$  has type  $(-1, 0)$  then  $S_1 = \sum_{i=0}^{p-1} (\mathbb{k}x_2^i \partial_1)$  and  $S_i = 0$  for  $i \geq 2$ . Moreover, the Lie algebra

$$S_0 = \sum_{i=0}^{p-1} (\mathbb{k}x_2^i \partial_2) \oplus \sum_{i=0}^{p-1} \mathbb{k}x_2^i (x_1 \partial_1)$$

is isomorphic to  $W(1; \underline{1}) \ltimes O(1; \underline{1})$  where  $\sum_{i=0}^{p-1} \mathbb{k}x_2^i (x_1 \partial_1) \cong O(1; \underline{1})$  is an abelian ideal of  $S_0$ . It is straightforward to check that  $\mathcal{G}_0 = S_0$  acts faithfully on  $\mathcal{G}_1 = S_1$ . So the equality  $\bar{Q} = 0$  must hold. Let  $\bar{R} = \mathbb{k}\partial_2 \oplus S_1$  and denote by  $R$  the inverse image of  $\bar{R}$  in  $M_{(0)}$ . Then  $[R, R] \subseteq M_{(1)}$  consists of nilpotent elements of  $\mathfrak{g}$ . In view of Corollary 2.3 we may assume that  $[R, e] \subseteq \mathbb{k}e$ . Then  $[\partial_2, \bar{e}] = 0$  forcing  $\bar{e} = \lambda \partial_1$  for some  $\lambda \in \mathbb{k}$ . This, however, contradicts the fact that  $(\text{ad } \partial_1)^4 \neq 0$ .

As a result, we may assume that the grading of  $S$  is either standard or has type  $(1, 0)$ . These are the reverse gradings of the ones considered earlier. In particular,  $\bar{\mathcal{G}}_{-2} = S_{-2} = 0$ . Let  $\mathfrak{t} = \mathbb{k}(x_1 \partial_1) \oplus \mathbb{k}(x_2 \partial_2)$ , a maximal toral subalgebra of  $\bar{\mathcal{G}} = S$  contained in  $\mathcal{G}_0 = S_0$ . Since  $\mathcal{G}_0$  is

restricted, there exists a 2-dimensional toral subalgebra  $\tilde{\mathfrak{t}}$  in  $M_{(0)}$  which maps onto  $\mathfrak{t}$  under the canonical homomorphism  $M_{(0)} \twoheadrightarrow \mathcal{G}_0$ ; see [Str04, Proposition 1.2.2]. Since  $M$  is non-regular and  $\mathfrak{c}_{\bar{\mathcal{G}}}(\mathfrak{t}) = \mathfrak{c}_S(\mathfrak{t}) = \mathfrak{t}$  it must be that  $N(\mathcal{G}) \neq 0$ . So the above yields  $0 \neq \mathcal{G}_{-2} \subseteq N(\mathcal{G})$ . Since  $\text{Der}(S) = \text{ad } S$ , the Lie algebra  $\mathcal{G}_0$  contains the degree derivation  $t_0$  associated with our present grading of  $S$ . As before, we find a toral element  $\tilde{t}_0 \in M_{(0)}$  which maps onto  $t_0$  under the canonical homomorphism  $M_{(0)} \twoheadrightarrow \mathcal{G}_0$  and pick, for any  $i \in \mathbb{Z}$ , a subspace  $V_i$  of  $\mathfrak{g}$  such that  $M_{(i)} = V_i \oplus M_{(i+1)}$  and  $[\tilde{t}_0, v] = \bar{i}v$  for all  $v \in V_i$ . Repeating verbatim the argument used in the  $W(1; \underline{n})$ -case we obtain that  $[V_{-2}, M_{(i)}] \subseteq M_{(1)}$  for  $i = 1, 2$ . It follows that  $0 \neq V_{-2} \subset \mathfrak{n}_{\mathfrak{g}}(\text{nil}(M)) = M$  which is impossible. This enables us to conclude that the case where  $S \cong W(r; \underline{d})$  cannot occur.

3.15. Suppose  $S$  is of special Cartan type. Then it follows from [Str04, 6.3] that  $S$  is one of

$$S(r; \underline{d})^{(1)}, \quad S(r; \underline{d}; \Phi(\tau))^{(1)}, \quad S(r; \underline{d}; \Phi(l))^{(1)},$$

where  $r \geq 3$  and  $1 \leq l \leq r$ . The Lie algebras in this list have dimensions  $(r-1)(p^{|\underline{d}|} - 1)$ ,  $(r-1)(p^{|\underline{d}|} - 1)$  and  $(r-1)p^{|\underline{d}|}$ , respectively. As a consequence,  $\dim S \geq 2(p^3 - 1) \geq 248$ . Since  $\dim \mathfrak{g} \geq \dim S$  and  $p > 5$  if  $\mathfrak{g}$  is of type  $E_8$ , we see that this case is impossible.

3.16. Suppose  $S$  is a Hamiltonian Lie algebra, that is  $S \cong H(2r; \underline{d}; \Phi)^{(2)}$ , where  $r \geq 1$ . This case is more complicated because several entirely different situations may occur here. We first note that  $\dim S \geq \dim H(2r; \underline{d})^{(2)}$  because  $H(2r; \underline{d}; \Phi)^{(2)}$  is a filtered deformation of graded Lie algebra sandwiched between the *graded* Hamiltonian algebras  $H(2r; \underline{d})^{(2)}$  and  $CH(2r; \underline{d})$ ; see [Str04, Theorem 6.1.2], for instance. Hence  $\dim S \geq \dim H(2r; \underline{1})^{(2)} = p^{2r} - 2$ . If  $r \geq 2$  then  $248 \geq \dim \mathfrak{g} \geq \dim S \geq 5^4 - 2$ , a contradiction. In view of [Str04, Theorem 6.3.10] this shows that  $S$  is one of

$$H(2; \underline{d})^{(2)}, \quad H(2; \underline{d}; \Phi(\tau))^{(1)}, \quad H(2; \underline{d}; \Phi(l)), \quad l = 1, 2,$$

where  $\underline{d} \in \{(1, 1), (2, 1), (1, 2)\}$ . Besides,  $H(2; (1, 2); \Phi)^{(2)} \cong H(2; (2, 1); \Phi)^{(2)}$  unless  $\Phi = \Phi(l)$ .

3.17. We first assume that  $A(\bar{\mathcal{G}}) = S \otimes O(m; \underline{n})$  and  $m \geq 1$ . Then  $\dim \mathcal{G} > (\dim S)p^{|\underline{n}|} \geq (p^2 - 2)p$ . From this it is immediate that  $p = 5$ ,  $O(m; \underline{n}) = O(1; \underline{1})$ ,  $\mathfrak{g}$  is of type  $E_7$ , and  $S$  is one of  $H(2; \underline{1})^{(2)}$ ,  $H(2; \underline{1}, \Phi(\tau))^{(1)}$ ,  $H(2; \underline{1}; \Phi(1))$ .

Let  $\mathcal{M}(S)$  denote the standard maximal subalgebra of the Cartan type Lie algebra  $S$ . By Kreknin's theorem,  $\mathcal{M}(S)$  is invariant under all automorphisms of  $S$ ; see [Str04, Theorem 4.2.6]. Since our grading of  $S$  is induced by the action of a 1-dimensional torus  $T_0 \subset \text{Aut}(S)$ , this implies that  $\mathcal{M}(S) = \bigoplus_{i \in \mathbb{Z}} (\mathcal{M}(S) \cap S_i)$ . Since  $\mathcal{M}(S)$  is maximal in  $S$ , it cannot consist of nilpotent elements of the minimal  $p$ -envelope  $S_p$  of  $S$ . Since each  $S_i \cap \mathcal{M}(S)$  with  $i \neq 0$  acts nilpotently on  $S$ , it follows from the Engel–Jacobson theorem that  $S_0 \cap \mathcal{M}(S)$  contains a non-nilpotent element of  $S_p$ . On the other hand, it is immediate from [Str04, Lemma 7.1.1.(3)] that  $\mathcal{M}(S)$  is a restricted subalgebra of  $S_p$ . It follows that so is  $\mathcal{M}(S) \cap S_0 = \mathcal{M}(S)^{T_0}$ . As a consequence,  $\mathcal{M}(S) \cap S_0$  contains a nonzero toral element,  $h_0$  say. It is straightforward to see that  $S_p \otimes 1$  coincides with the  $p$ -envelope of  $S \otimes 1$  in  $\text{Der}(S \otimes O(1; \underline{1}))$ .

Let  $C_0 = \mathfrak{c}_{\bar{\mathcal{G}}_0}(h_0 \otimes 1)$ . Clearly,  $h_0 \otimes O(1; \underline{1}) \subset C_0$ . Recall that

$$S_0 \otimes O(1; \underline{1}) \subset \bar{\mathcal{G}}_0 \subseteq (\text{Der}_0(S) \otimes O(1; \underline{1})) \rtimes (\text{Id}_S \otimes \mathcal{D})$$

for some transitive Lie subalgebra  $\mathcal{D}$  of  $W(1; \underline{1})$ . Let  $\mathfrak{m} = O(1; \underline{1})x_1$  denote the maximal ideal of  $O(1; \underline{1}) = O(x_1)$ . Since  $\text{ad}_{\bar{\mathcal{G}}}(h_0 \otimes 1)$  is semisimple and  $\mathcal{D}$  is transitive,  $C_0$  contains an element  $c = (d \otimes f) + (\text{Id}_S \otimes D)$  with  $d \in \text{Der}_0(S)$ ,  $f \in O(1; \underline{1})$ ,  $D \in W(1; \underline{1})$  such that  $D(x_1) \notin \mathfrak{m}$ . Since

$c \in C_0$  it must be that  $[d, h_0] = 0$ . By construction,  $[c, h_0 \otimes x_1] = h_0 \otimes D(x_1)$ . Since  $h_0^{[p]} = h_0$  and  $D(x_1)^p$  is a *nonzero* scalar in  $O(1; \underline{1})$ , the  $p$ -closure of  $[c, h_0 \otimes x_1]$  in the restricted Lie algebra  $\bar{\mathcal{G}}_0$  contains  $h_0 \otimes 1$ .

Since the canonical homomorphism  $\eta: M_{(0)} \rightarrow \bar{\mathcal{G}}_0$  is restricted, there is a toral element  $\tilde{h}_0 \in M_{(0)}$  such that  $\eta(\tilde{h}_0) = h_0 \otimes 1$ . Let  $\tilde{C}_0 = \mathfrak{c}_M(\tilde{h}_0)$ . Since  $\text{ad } \tilde{h}_0$  is semisimple, we have that  $\eta(\tilde{C}_0) = C_0$  and  $\tilde{C}_0 \cap \ker \eta = \mathfrak{c}_{M_{(1)}}(\tilde{h}_0)$ . Since both  $c$  and  $h_0 \otimes x_1$  are in  $C_0$ , the above discussion implies that the  $p$ -closure of  $[\tilde{C}_0, \tilde{C}_0]$  in  $M$  contains an element of the form  $\tilde{h}_0 + y$  for some  $y \in \tilde{C}_0 \cap M_{(1)}$ . As  $(\tilde{h}_0 + y)^{[p]^N} = \tilde{h}_0^{[p]^N} = \tilde{h}_0$  for  $N \gg 0$ , it follows that the  $p$ -closure of the derived subalgebra of  $\mathfrak{g}_{\tilde{h}_0}$  contains  $\tilde{h}_0$ . Since it follows from Jacobson's formula that for any Levi subalgebra  $\mathfrak{l}$  of  $\mathfrak{g}$  the Lie algebra  $[\mathfrak{l}, \mathfrak{l}]$  is closed under taking  $p$ -th powers in  $\mathfrak{g}$ , this yields  $\tilde{h}_0 \in [\mathfrak{g}_{\tilde{h}_0}, \mathfrak{g}_{\tilde{h}_0}]$ .

We now wish to estimate  $\dim \mathfrak{g}(\tilde{h}_0, i)$  for all  $i \in \mathbb{F}_p^\times$ . First we note that since  $\text{ad } \tilde{h}_0$  is toral and preserves all components  $M_{(j)}$  of our filtration, the equality  $\dim \mathfrak{g}(\tilde{h}_0, i) = \dim \mathcal{G}(h_0 \otimes 1, i)$  must hold for all  $i \in \mathbb{F}_p$ . Next we observe that for any  $k \geq 1$  the factor space  $N(\mathcal{G})^k / N(\mathcal{G})^{k+1}$  is a graded  $\bar{\mathcal{G}}$ -module. This yields

$$(10) \quad \dim \mathfrak{g}(\tilde{h}_0, i) = \dim \bar{\mathcal{G}}(h_0 \otimes 1, i) + \sum_{k \geq 1} \dim (N(\mathcal{G})^k / N(\mathcal{G})^{k+1})(h_0 \otimes 1, i).$$

Clearly, each  $N(\mathcal{G})^k / N(\mathcal{G})^{k+1}$  is a graded  $\bar{\mathcal{G}}$ -module. Since  $\bar{\mathcal{G}}$  is semisimple and  $\mathcal{G}$  is a restricted  $(\text{ad } \mathcal{G}_0)$ -module, applying Lemma 2.5 shows that all composition factors of the  $(S \otimes 1)$ -modules  $\bar{\mathcal{G}}$  and  $N(\mathcal{G})^k / N(\mathcal{G})^{k+1}$  with  $k \geq 1$  are restrictable.

As  $S \in \{H(2; \underline{1})^{(2)}, H(2; \underline{1}; \Phi(\tau))^{(1)}, H(2; \underline{1}; \Phi(1))\}$ , any non-trivial irreducible restricted  $S_p$ -module  $V$  is either isomorphic to  $S$  (the adjoint  $(\text{ad } S)$ -module) or is induced from an irreducible restricted  $\mathcal{M}(S)$ -module  $V_0$ ; see [FSW15, § 5] or [HS15b, Lemma 2.7]. Note that  $\mathcal{M}(S)/\text{nil}(\mathcal{M}(S))$  is isomorphic to  $\mathfrak{sl}_2$  or  $\mathfrak{gl}_2$  and  $S/\mathcal{M}(S)$  is a 2-dimensional irreducible module over  $\mathcal{M}(S)/\text{nil}(\mathcal{M}(S))$ . Since  $h_0 \in \mathcal{M}(S)$  this implies that if  $V$  is induced from  $V_0$  then  $\dim V(h_0, i) = p(\dim V_0)$  for all  $i \in \mathbb{F}_p$ . If  $V \cong S$  then one derives that  $\dim V(h_0, i) = p$  for all  $i \in \mathbb{F}_p^\times$  by passing from the filtered Cartan type algebra  $S$  to the corresponding graded Lie algebra  $\text{gr}(S) \supseteq H(2; \underline{1})^{(2)}$ . In conjunction with (10) this entails that  $\dim \mathfrak{g}(\tilde{h}_0, i) = \dim \mathfrak{g}(\tilde{h}_0, j)$  for all  $i, j \in \mathbb{F}_p^\times$ .

As a result,  $\tilde{h}_0$  is a  $p$ -balanced toral element of  $\mathfrak{g}$ . Since in the present case  $p = 5$  and  $G$  is of type  $E_7$ , applying Proposition 2.7(ii) yields that the Levi subalgebra  $\mathfrak{g}_{\tilde{h}_0}$  has type  $D_4A_1$ . But then the derived subalgebra of  $\mathfrak{g}_{\tilde{h}_0}$  is semisimple and hence cannot contain its central element  $\tilde{h}_0$ . This contradiction shows that the case where  $m \geq 1$  cannot occur.

3.18. In the next five subsections we assume that  $A(\bar{\mathcal{G}}) = S \cong H(2; \underline{1})^{(2)}$ . In this case we may regard  $\bar{\mathcal{G}}$  as a graded subalgebra of  $\text{Der}(H(2; \underline{1})^{(2)})$ . Recall that  $S$  is spanned by all

$$D_H(f) := \partial_1(f)\partial_2 - \partial_2(f)\partial_1$$

where  $f \in O(2; \underline{1})$  is a linear combination of  $x_1^{a_1}x_2^{a_2}$  with  $0 < a_1 + a_2 < 2(p-1)$ , and the standard maximal subalgebra  $\mathcal{M}(S)$  has a basis consisting of all  $D_H(x_1^{a_1}x_2^{a_2})$  with  $2 \leq a_1 + a_2 < 2(p-1)$ . By [PS99, Theorem 3.3], every grading of  $\text{Der}(H(2; \underline{1})^{(2)})$  is induced by a suitable  $(a_1, a_2)$ -grading of  $W(2; \underline{1})$  (which contains the derivation algebra of  $H(2; \underline{1})^{(2)}$ ). Since  $\bar{\mathcal{G}}_{-1}$  is an irreducible  $\mathcal{G}_0$ -module, no generality will be lost by assuming that  $(a_1, a_2) \in \{(-1, -1), (-1, 0), (1, 0), (1, 1)\}$ ; see [PS99, Corollary 3.4(4)]. If  $a_1 = a_2 = -1$  then  $\mathcal{G}_i = 0$  for  $i \geq 2$ ,  $\bar{\mathcal{G}}_1 = \mathbb{k}D_H(x_1) \oplus \mathbb{k}D_H(x_2) = \mathbb{k}\partial_2 \oplus \mathbb{k}\partial_1$ ,

and

$$\mathfrak{sl}_2 \cong S_0 \subseteq \bar{\mathcal{G}}_0 \subseteq \mathfrak{gl}_2$$

acts irreducibly on  $\bar{\mathcal{G}}_1$ . From this it is immediate that  $M_{(1)} \cong \bar{\mathcal{G}}_1$  and  $\mathfrak{c}_{\mathfrak{g}}(M_{(1)})$  coincides with  $M_{(1)}$  (recall that  $\mathfrak{c}_{\mathfrak{g}}(M_{(1)}) \subseteq M_{(0)}$  by the maximality of  $M$ ). Since  $p \geq 5$ , every nonzero element  $D \in \mathbb{k}\partial_1 \oplus \mathbb{k}\partial_2$  has the property that  $(\text{ad } D)^4(\bar{\mathcal{G}}) \neq 0$ . Arguing as before it is now easy to observe that this contradicts Corollary 2.3. So the case where  $(a_1, a_2) = (-1, -1)$  is impossible.

3.19. If  $(a_1, a_2) = (-1, 0)$  then [PS99, Corollary 3.4] implies that  $\bar{\mathcal{G}}_i = 0$  for  $i \geq 2$  and  $\bar{\mathcal{G}}_0$  is sandwiched between  $S_0 = \bigoplus_{i=0}^{p-1} \mathbb{k}(ix_2^{i-1}x_1\partial_1 - x_2^i\partial_2) \cong W(1; \underline{1})$  and  $S_0 \oplus \mathbb{k}z$  where  $z = x_1\partial_1$  acts on  $\bar{\mathcal{G}}_1$  as a scalar operator. Also,  $S_1 = \bigoplus_{i=0}^{p-2} \mathbb{k}x_2^i\partial_1 \cong O(1; \underline{1})/\mathbb{k}1$  has codimension  $\leq 1$  in  $\bar{\mathcal{G}}_1 \subseteq S_1 \oplus \mathbb{k}x_2^{p-1}\partial_1$  which is indecomposable as a  $\mathbb{k}\partial_2$ -module. We now see that as in the previous case  $M_{(1)} \cong \bar{\mathcal{G}}_1$  is abelian and  $\mathfrak{c}_{\mathfrak{g}}(M_{(1)}) = M_{(1)}$ . We pick an element  $v \in M_{(0)}$  which maps onto  $\partial_2 \in S_0$  under the canonical homomorphism  $M_{(0)} \rightarrow \mathcal{G}_{(0)} = M_{(0)}/M_{(1)}$  and set  $R = \mathbb{k}v$ . Since  $\mathfrak{c}_{\bar{\mathcal{G}}_1}(\partial_2) = \mathbb{k}\partial_1$  it follows from Corollary 2.3 that there exists an element  $e \in \mathcal{O}_{\min} \cap M_{(1)}$  which maps onto  $\partial_1$  under the epimorphism  $M_{(1)} \twoheadrightarrow \mathcal{G}_{(1)}$ . As  $(\text{ad } e)^3 = 0$  this contradicts the fact that  $(\text{ad } \partial_2)^4(H(2; \underline{1})^{(2)}) \neq 0$ . Hence the case where  $(a_1, a_2) = (-1, 0)$  cannot occur either.

3.20. Now suppose  $(a_1, a_2) = (1, 0)$ . Then [PS99, Corollary 3.4(2)] implies that  $S = \bigoplus_{i=-1}^{p-2} S_i$  and  $\bar{\mathcal{G}} = \bigoplus_{i=-1}^r \bar{\mathcal{G}}_i$  where  $r = p - 2$  or  $r = p - 1$ . Moreover, if  $r = p - 1$  then  $\dim \mathcal{G}_r = \dim \bar{\mathcal{G}}_r = 1$ . Since  $\dim \mathcal{G}_r = \dim M_{(r)}$  the equality  $r = p - 1$  would mean that  $M$  normalises a 1-dimensional subspace  $\mathbb{k}e \subset \mathcal{N}(\mathfrak{g})$ . As the latter would entail that  $M \subseteq \mathfrak{n}_{\mathfrak{g}}(\mathbb{k}e)$  lies in the optimal parabolic subalgebra of  $e$ , it must be that  $r = p - 2$ .

Since  $\bar{\mathcal{G}}$  is a graded Lie algebra, the centre of  $\bar{\mathcal{G}}_{>0}$  is a graded subalgebra of  $\bar{\mathcal{G}}$ . Since  $S \cong \text{ad } S$  coincides the unique minimal ideal of  $\bar{\mathcal{G}} \subseteq \text{Der}(S)$  and  $S_{p-2} \neq 0$ , this forces  $\mathfrak{z}(\bar{\mathcal{G}}_{>0}) = \bar{\mathcal{G}}_{p-2}$ . From this it is immediate that  $\mathfrak{c}_{\mathfrak{g}}(\text{nil}(M)) = \mathfrak{z}(M_{(1)}) = M_{(p-2)}$ .

Applying [PS99, Corollary 3.4(2)] it is straightforward to see that the endomorphism  $\text{ad } D_H(x_1) = \partial_2 \in \text{ad } S_0$  acts on  $\mathcal{G}_{p-2} = \bar{\mathcal{G}}_{p-2}$  as a single Jordan block and annihilates  $D_H(x_1^{p-1}) = -x_1^{p-2}\partial_2 \in S_{p-2}$ . Pick  $e \in M_{(p-2)}$  which maps onto  $D_H(x_1^{p-1})$  under the canonical epimorphism  $M_{(p-2)} \twoheadrightarrow \bar{\mathcal{G}}_{p-2}$ . Using Corollary 2.3 and arguing as in (3.19) we now observe that  $e \in \mathcal{O}_{\min}$ .

Let  $E$  be a non-trivial restricted irreducible  $S$ -module. We have already mentioned in (3.17) that either  $E \cong S$ , the adjoint  $S$ -module, or  $E$  is induced from an irreducible restricted module  $E_0$  over  $\mathcal{M}(S)/\text{nil}(\mathcal{M}(S)) \cong \mathfrak{sl}_2$ . Let  $\rho: S \rightarrow \mathfrak{gl}(E)$  denote the corresponding representation. If  $\rho = \text{ad}$  then Since  $\rho(D_H(x_1^{p-1}))$  is injective on the span of  $\{D_H(x_2^i), D_H(x_1x_2^i) \mid 1 \leq i \leq p-1\}$  implying that  $\dim \text{Im } \rho(D_H(x_1^{p-1})) \geq 2(p-1)$ . Now suppose  $E \cong u(S) \otimes_{u(\mathcal{M}(S))} E_0$  and write  $x \cdot v$  for  $(\rho(x))(v)$  where  $x \in S$  and  $v \in E$ . Let  $w \in E$  be such that  $D_H(x_1^k) \cdot w = 0$  for  $3 \leq k \leq p-1$ . Recall that  $D_H(x_1^k) = kx_1^{k-1}\partial_2$  and  $x_1\partial_2 = \frac{1}{2}D_H(x_1^2) \in \mathcal{M}(S)$  acts nilpotently on the subspace  $E_0$  of  $E$  and commutes with  $\partial_2$ .

We have  $(x_1\partial_2) \cdot \partial_1 \cdot w = -\partial_2 \cdot w + \partial_1 \cdot (x_1\partial_2)(w)$ . Suppose

$$(11) \quad (x_1^k\partial_2) \cdot \partial_1^k \cdot w = k!(-1)^k\partial_2 \cdot w + \lambda_k\partial_1 \cdot (x_1\partial_2)(w) \quad \text{for some } \lambda_k \in \mathbb{k}.$$



Then

$$\begin{aligned}
(x_1^{k+1}\partial_2) \cdot \partial_1^{k+1} \cdot w &= [x_1^{k+1}\partial_2, \partial_1] \cdot \partial_1^k \cdot w + \partial_1 \cdot (x_1^{k+1}\partial_2) \cdot \partial_1^k \cdot w \\
&= -(k+1)(x_1^k\partial_2) \cdot \partial_1^k \cdot w + (k+1)!(-1)^k \partial_1 \cdot (x_1\partial_2)(w) \\
&= (k+1)!(-1)^{k+1}\partial_2 \cdot w + ((k+1)!(-1)^k - (k+1)\lambda_k) \partial_1 \cdot (x_1\partial_2)(w).
\end{aligned}$$

This shows that (11) holds for all  $k \in \{1, \dots, p-2\}$ . Since  $(p-2)! = 1$  in  $\mathbb{k}$  we get

$$(12) \quad D_H(x_1^{p-1}) \cdot \partial_1^{p-2} \cdot w = \partial_2 \cdot w - \lambda_{p-2} \partial_1 \cdot (x_1\partial_2)(w).$$

Next we show that the equality

$$(13) \quad (x_1^k\partial_2) \cdot \partial_1^{k+1} \cdot w = (k+1)!(-1)^k \partial_1\partial_2 \cdot w + \mu_k \partial_1^2 \cdot (x_1\partial_2)(w) \quad \text{with } \mu_k \in \mathbb{k}$$

holds for all  $k \in \{1, \dots, p-2\}$ . Since

$$(x_1\partial_2) \cdot \partial_1^2 \cdot w = [x_1\partial_2, \partial_1] \cdot \partial_1 \cdot w + \partial_1 \cdot [x_1\partial_2, \partial_1] \cdot w + \partial_1^2 \cdot (x_1\partial_2)(w) = -2\partial_1\partial_2 \cdot w + \partial_1^2 \cdot (x_1\partial_2)(w)$$

the statement holds for  $k = 1$ . If it holds for  $k = m$  then using (11) we get

$$\begin{aligned}
(x_1^{m+1}\partial_2) \cdot \partial_1^{m+2} \cdot w &= \partial_1 \cdot (x_1^{m+1}\partial_2) \cdot \partial_1^{m+1} \cdot w + [x_1^{m+1}\partial_2, \partial_1] \cdot \partial_1^{m+1} \cdot w \\
&= (m+1)!(-1)^{m+1} \partial_1\partial_2 \cdot w + \lambda_m \partial_1^2 \cdot (x_1\partial_2)(w) - (m+1)(x_1^m\partial_2) \cdot \partial_1^{m+1} \cdot w \\
&= ((m+1)!(-1)^{m+1} - (m+1)!(-1)^m(m+1)) \cdot \partial_1\partial_2 \cdot w \\
&\quad + (\lambda_m - (m+1)\mu_m) \partial_1^2 \cdot (x_1\partial_2)(w) \\
&= (m+2)!(-1)^{m+1} \partial_1\partial_2 \cdot w + \mu_{m+1} \partial_1^2 \cdot (x_1\partial_2)(w).
\end{aligned}$$

This proves (13). As a consequence,

$$(14) \quad D_H(x_1^{p-1}) \cdot \partial_1^{p-1} \cdot w = -\partial_1\partial_2 \cdot w - \mu_{p-2} \partial_1^2 \cdot (x_1\partial_2)(w).$$

Note that we can take for  $w$  any vector of the form  $\partial_2^i \otimes v$  with  $v \in E_0$  and  $0 \leq i \leq p-2$ . Applying (12) and (14) it is now straightforward to see that  $\rho(D_H(x_1^{p-1}))$  is injective on the subspace  $\bigoplus_{i=0}^{p-2} ((\partial_1^{p-1}\partial_2^i \otimes E_0) \oplus (\partial_1^{p-2}\partial_2^i \otimes E_0))$  of  $E$ . It follows that  $\text{Im } \rho(D_H(x_1^{p-1}))$  has dimension  $\geq 2(p-1)(\dim E)/p^2$ .

Let  $\text{gr}(\mathcal{G})$  denote the  $S$ -module  $\bar{\mathcal{G}} \oplus \sum_{k \geq 1} N(\mathcal{G})^k / N(\mathcal{G})^{k+1}$  and write  $\varphi$  for the corresponding representation of  $S$ . Let  $E_1, \dots, E_r$  be all induced composition factors of the  $S$ -module  $\text{gr}(\mathcal{G})$  and write  $l$  for the multiplicity of  $S$  in  $\text{gr}(\mathcal{G})$ . Put

$$s := l + \sum_{i=1}^r (\dim E_i)/p^2.$$

Repeating almost verbatim the argument used in (3.17) one observes that  $\mathfrak{g}$  contains a  $p$ -balanced toral element  $\tilde{h}_0$  such that  $\dim \mathfrak{g}(\tilde{h}_0, i) = ps$  for all  $i \in \mathbb{F}_p^\times$ . On the other hand, since  $D_H(x_1^{p-1}) = \text{gr}(e)$  for some  $e \in M_{(p-2)} \cap \mathcal{O}_{\min}$ , the above discussion yields

$$(15) \quad \dim \mathcal{O}_{\min} = \dim (\text{Im ad } e) \geq \dim (\text{Im } \varphi(D_H(x_1^{p-1}))) \geq 2(p-1)s.$$

If  $G$  is of type  $E_6$  then Proposition 2.7(i) shows that  $p = 5$  and  $s = 3$ . But then  $2(p-1)s = 24 > 22 = \dim \mathcal{O}_{\min}$ . As this contradicts (15) this case cannot occur. If  $\mathfrak{g}$  is of type  $E_7$  then Proposition 2.7(ii) yields that  $p = 5$  and  $s = 5$ . Since  $\dim \mathcal{O}_{\min} = 34 < 40 = 2(p-1)s$  this case cannot occur either.

3.21. Suppose  $G$  is of type  $E_8$ . Then Proposition 2.7 says that  $p = 7$  or  $p = 11$ . Moreover, if  $p = 7$  then  $\mathfrak{g}_{\tilde{h}_0}$  has dimension 80 or 38. If  $\dim \mathfrak{g}_{\tilde{h}_0} = 38$  then  $s = 5$  and  $\dim \mathcal{O}_{\min} = 58 < 60 = 2(p-1)s$  which violates (15).

Suppose  $p = 7$  and  $\dim \mathfrak{g}_{\tilde{h}_0} = 80$ . Then  $s = 4$  and part (f) of the proof of Proposition 2.7 shows that no generality will be lost by assuming that  $\tilde{h}_0 = 2t_7 + 2t_8$ . In this case, there exists an element  $e \in \mathcal{O}(D_4)$  which admits an optimal cocharacter  $\lambda: \mathbb{k}^\times \rightarrow G$  with the property that  $e \in \mathfrak{g}(\lambda_e, 2)$  and  $(d\lambda)(1) = \tilde{h}_0$ . Moreover, it follows from (7) and the preceding discussion that the weight space  $\mathfrak{g}(\lambda, 2)$  coincides with the 2-eigenspace of  $\text{ad } \tilde{h}_0$ . In view of [Pre03, Theorem 2.3] this implies that the subset  $\mathfrak{g}(\tilde{h}_0, 2) \cap \mathcal{O}(D_4)$  is Zariski dense in  $\mathfrak{g}(\tilde{h}_0, 2)$ . As  $\mathcal{O}(D_4) \subset \mathcal{N}_p(\mathfrak{g})$  by [Law95, Table 9] and [PSt15, Theorem 4.1] we now deduce that  $x^{[p]} = 0$  for all  $x \in \mathfrak{g}(\tilde{h}_0, 2)$ . We may assume that the image of  $\tilde{h}_0$  in  $\mathcal{G}_0$  equals  $2D_H(x_1x_2)$ .

As  $[2D_H(x_1x_2), D_H(x_2)] = 2D_H(x_2)$  we can find an element  $u \in \mathfrak{g}(\tilde{h}_0, 2) \cap M_{(0)}$  which maps onto  $D_H(x_2) = -\partial_1$  under the epimorphism  $M_{(0)} \twoheadrightarrow \mathcal{G}_0$ . One checks directly that  $\text{Im}(\text{ad}_S D_H(x_2)^{p-1})$  has dimension  $p-1$ . Since  $D_H(x_2) \notin S_{(0)}$ , it is straightforward to see for any irreducible restricted representation  $\rho: S \rightarrow \mathfrak{gl}(E)$  such that  $E = u(S) \otimes_{u(S_{(0)})} E_0$  as  $S$ -modules,  $\rho(D_H(x_2))$  has  $p \dim E_0$  Jordan blocks of size  $p$ . Since  $u^{[p]} = 0$  by our earlier remark, the definition of  $s$  in (3.20) shows that  $\text{ad } u$  has at least  $4(p-1) = 24$  Jordan blocks of size  $p = 7$ .

As  $\dim(\text{Ad } G)u \leq \dim(\text{Ad } G)e = 168$ , combining [Law95, Table 9] with [PSt15, Theorem 4.1] now yields  $u \in \mathcal{O}(D_4)$ . Note that  $D_H(x_1^k x_2^{p-1}) \in S_{k-1}$  for  $1 \leq k \leq p-2$ . Since  $(\text{ad } u)^p = 0$  and

$$(\text{ad } D_H(x_2))^{p-1}(D_H(x_1^k x_2^{p-1})) \neq 0, \quad [2D_H(x_1x_2), D_H(x_1^k x_2^{p-1})] = -2(k+1)D_H(x_1^k x_2^{p-1}),$$

it follows that  $\mathfrak{g}_u \cap \mathfrak{g}(\tilde{h}_0, 2j) \neq 0$  for five values of  $j \in \mathbb{F}_p^\times$ . This, however, contradicts [LT11, p. 131]. Therefore, the case where  $p = 7$  cannot occur.

If  $p = 11$  then  $[M_{(5)}, M_{(5)}] \subseteq M_{(10)} = M_{(p-1)} = 0$  and

$$\dim M_{(5)} = \sum_{i=5}^9 \mathcal{G}_i \geq \dim \sum_{i=5}^9 S_i = 4p + (p-1) = 54.$$

Hence  $\mathfrak{g}$  contains an abelian Lie subalgebra of dimension 54. The set of all  $d$ -dimensional abelian Lie subalgebras of  $\mathfrak{g}$  is a closed subset of the Grassmannian  $\text{Gr}(d, \mathfrak{g})$  invariant under the action of a maximal torus  $T$  of  $G$ . So it follows from the Borel fixed-point theorem that the root system  $\Phi = \Phi(G, T)$  contains an *abelian* subset  $\mathcal{A}$  of size 54 (it has the property that  $\beta + \gamma \notin \Phi$  for all  $\beta, \gamma \in \mathcal{A}$ ). However, an old result of Malcev says that in type  $E_8$  the size of any abelian subset of  $\Phi$  cannot be bigger than 36; see [Ma45]. It is not hard to see that this result of Malcev is still valid in our situation; see [PeS15] for detail. Thus  $G$  is not of type  $E_8$ .

If  $G$  is of type  $F_4$  then  $p = 7$  by Proposition 2.7. Hence  $[M_{(3)}, M_{(3)}] \subseteq M_{(6)} = M_{(p-1)} = 0$ . Since

$$\dim M_{(3)} = \sum_{i=3}^5 \mathcal{G}_i \geq \dim \sum_{i=3}^5 S_i = 2p + (p-1) = 20.$$

Arguing as before we now deduce that a root system of type  $F_4$  contains an abelian subset of size 20. Since such subsets cannot have size bigger than 9 by [Ma45] and [PeS15], we conclude that the case where  $(a_1, a_2) = (1, 0)$  is impossible.

3.22. Finally, suppose  $(a_1, a_2) = (1, 1)$ , i.e. the grading of the Cartan type Lie algebra  $S = H(2; \underline{1})^{(2)}$  is standard. Then  $S_0 = \mathbb{k}D_H(x_1^2) \oplus \mathbb{k}D_H(x_1x_2) \oplus \mathbb{k}D_H(x_2^2)$  is isomorphic to  $\mathfrak{sl}_2$  and  $\bar{G}_{-1} = S_{-1} = \mathbb{k}D_H(x_1) \oplus \mathbb{k}D_H(x_2)$  is a 2-dimensional irreducible  $S_0$ -module. Furthermore,  $\bar{G}_k = 0$

for  $k \leq -2$  and  $\mathcal{G}_1 = S_1 = \mathbb{k}D_H(x_1^3) \oplus \mathbb{k}D_H(x_1^2x_2) \oplus \mathbb{k}D_H(x_1x_2^2) \oplus \mathbb{k}D_H(x_2^3)$  is an irreducible 4-dimensional  $S_0$ -module.

As before, we choose a toral element  $\tilde{h}_0 \in M_{(0)}$  which maps onto  $D_H(x_1x_2)$  under the epimorphism  $M_{(0)} \twoheadrightarrow \mathcal{G}_0$ . Since the endomorphism  $\text{ad } \tilde{h}_0$  is semisimple, there exist  $v_{-1,\pm 1} \in M_{(-1)} \cap \mathfrak{g}(\tilde{h}_0, \pm 1)$  and  $v_{0,\pm 2} \in M_{(0)} \cap \mathfrak{g}(\tilde{h}, \pm 2)$  such that

$$\text{gr}_{-1}(v_{-1,-1}) = D_H(x_1), \quad \text{gr}_{-1}(v_{-1,1}) = D_H(x_2), \quad \text{gr}_0(v_{0,-2}) = D_H(x_1^2), \quad \text{gr}_0(v_{0,2}) = D_H(x_2^2).$$

Besides, there exist  $v_{1,\pm 3} \in M_{(1)} \cap \mathfrak{g}(\tilde{h}, \pm 3)$  and  $v_{1,\pm 1} \in M_{(1)} \cap \mathfrak{g}(\tilde{h}, \pm 1)$  such that

$$\text{gr}_1(v_{1,-3}) = D_H(x_1^3), \quad \text{gr}_1(v_{1,-1}) = D_H(x_1x_2^2), \quad \text{gr}_1(v_{1,1}) = D_H(x_1^2x_2), \quad \text{gr}_1(v_{1,3}) = D_H(x_2^3).$$

Since  $p > 3$  we have that  $\mathcal{G}_2 = S_2 = [S_1, S_1]$ . Our earlier remarks show that

$$(16) \quad [v_{0,-2}, v_{1,3}] \equiv b v_{1,1} \pmod{M_{(2)}} \quad \text{for some } b \in \mathbb{k}^\times.$$

To unify notation we put  $\tilde{h}_0 := v_{0,0}$ . If  $N(\mathcal{G}) = 0$  then  $\mathfrak{g} = M_{(-1)}$  contains  $\mathfrak{g}_{v_{0,0}}$ . Since the latter contains a maximal torus of  $\mathfrak{g}$  the subalgebra  $M$  is regular in  $\mathfrak{g}$ . So from now on we may assume that  $N_2(\mathcal{G}) = [\mathcal{G}_{-1}, \mathcal{G}_{-1}] \neq 0$ . In this case  $[\mathcal{G}_{-1}, \mathcal{G}_{-1}] \cong \wedge^2 \bar{\mathcal{G}}_{-1}$  is a 1-dimensional  $S_0$ -module and  $M_{(-2)} = [M_{(-1)}, M_{(-1)}] + M_{(-1)} = \mathbb{k}w \oplus M_{(-1)}$  where  $w = [v_{-1,-1}, v_{-1,1}] \neq 0$ .

We claim that  $w \in \mathfrak{n}_{\mathfrak{g}}(M_{(1)})$ . Since  $w \in M_{(-2)}$  and  $M_{(2)} = [M_{(1)}, M_{(1)}] + M_{(3)}$  we just need to show that  $[w, v_{i,1}] \in M_{(1)}$  for  $i \in \{\pm 1, \pm 3\}$ . Since  $w \in M_{(-2)} \cap \mathfrak{g}_{v_{0,0}}$  and

$$M_{(-2)} \subseteq \mathfrak{g}_{v_{0,0}} + \mathfrak{g}(v_{0,0}, 1) + \mathfrak{g}(v_{0,0}, -1) + M_{(0)}$$

we have  $[w, v_{0,\pm 2}] \in M_{(0)}$ . As  $[w, v_{1,3}] \in \mathfrak{g}(v_{0,0}, 3) \cap M_{(-1)}$  and  $M_{(-1)} \subseteq \mathfrak{g}(v_{0,0}, 1) + \mathfrak{g}(v_{0,0}, -1) + M_{(0)}$  it must be that  $[w, v_{1,3}] \in M_{(0)}$ . Hence  $[v_{0,-2}, [w, v_{1,3}]] \in M_{(0)}$ . But then (16) yields

$$\begin{aligned} M_{(0)} &\ni [v_{0,-2}, [w, v_{1,3}]] = [[v_{0,-2}, w], v_{1,3}] + [w, [v_{0,-2}, v_{1,3}]] \\ &\in [M_{(0)}, M_{(1)}] + b[w, v_{1,1}] + [w, M_{(2)}] \subseteq b[w, v_{1,1}] + M_{(0)}. \end{aligned}$$

As  $b \neq 0$  and  $M_{(0)} \cap \mathfrak{g}(v_{0,0}, 1) \subseteq M_{(1)}$  this yields  $[w, v_{1,1}] \in M_{(1)}$ . So  $[w, [v_{0,\pm 2}, v_{1,1}]] \in [M_{(0)}, v_{1,1}] + [v_{0,\pm 2}, M_{(1)}] \subseteq M_{(1)}$  forcing  $[w, v_{1,3}], [w, v_{1,1}] \in M_{(1)}$ . Finally,  $[w, [v_{0,-2}, v_{1,-1}]] \in [M_{(0)}, v_{1,-1}] + [v_{0,-2}, M_{(1)}] \subseteq M_{(1)}$ . Hence  $[w, v_{1,-3}] \in M_{(1)}$  proving the claim. As a result,  $w \in \mathfrak{n}_{\mathfrak{g}}(M_{(1)})$  contrary to the maximality of  $M$ . We now conclude that the case where  $S = H(2; \underline{1})^{(2)}$  is impossible.

3.23. Suppose that  $S \cong H(2; \underline{1}; \Phi)^{(1)}$  where  $\Phi$  is one of  $\Phi(\tau)$ ,  $\Phi(1)$ ,  $\Phi(2)$ . Then  $\text{Der}(S) = S_p$ ; see [Str04, Theorem 7.2.2]. The grading of  $S$  gives rise to that of  $\text{Der}(S)$ . Identifying  $S$  with  $\text{ad } S$  we have that  $S_i^{[p]} \subseteq \text{Der}_i(S)$  for all  $i \in \mathbb{Z}$ . Then it follows from Jacobson's formula for  $p$ -th powers that

$$(17) \quad \text{Der}_0(S) = \left( \sum_{i \in \mathbb{Z}, j \geq 0} S_i^{[p]^j} \right)_0 = \sum_{i \geq 0} S_0^{[p]^i}.$$

As  $\mathcal{G}_0 \cong \bar{\mathcal{G}}_0 \subseteq \text{Der}_0(S)$  this implies that  $S_{-1} = \mathcal{G}_{-1}$  is an irreducible  $S_0$ -module.

If  $\Phi = \Phi(\tau)$  then  $\omega = \Phi(\omega_S) = (1 + x_1^{p-1}x_2^{p-1})dx_1 \wedge dx_2$  and  $S$  is contained in the  $\mathbb{k}$ -span of all  $D_{H,\omega}(f)$  with  $f \in O(2; \underline{1})$ ; see [Str04, Theorem 6.5.7(2)]. Thanks to (the proof of) Lemma 2.8 we may assume that the grading of  $S$  is induced by the action of the torus  $T_\Phi = \{g(t, t^{-1}) \mid t \in \mathbb{k}^\times\}$ . Then  $S_0$  is contained in the  $\mathbb{k}$ -span of all  $D_{H,\omega}(x_1^k x_2^k)$  with  $k \in \{1, \dots, p-1\}$ . Using [Str04, (6.5.5)] it is then straightforward to check that  $S_0$  is abelian. The irreducibility of the  $S_0$ -module  $S_{-1}$  now gives  $\dim \mathcal{G}_{-1} = \dim S_{-1} = 1$ . But then  $\mathcal{G}_{-1}$  is not a faithful  $\mathcal{G}_0$ -module. This contradiction shows that  $S \not\cong H(2; \underline{1}; \Phi(\tau))^{(1)}$ .

Suppose  $\Phi = \Phi(l)$ . Since in the present case  $\underline{n} = \underline{1}$  we may assume that  $l = 1$ . Then  $\omega = \Phi(\omega_S) = \exp(x_1^{(p)})dx_1 \wedge dx_2$ . Using (17) we observe that  $\mathcal{G}_0 \cong \bar{\mathcal{G}}_0$  identifies with the  $p$ -closure of  $S_0$  in  $\text{Der}_0(S)$ . By the proof of Lemma 2.8, we may assume that the grading of  $S$  is induced by the action of the torus  $T_\Phi = \{g(1, t) \mid t \in \mathbb{k}^\times\}$ . In this case,  $S_0$  contains  $D_{H,\omega}(x_2) = -\partial_1 - x_1^{p-1}x_2\partial_2$ . Since  $\bar{\mathcal{G}}_0$  is a restricted subalgebra of  $\text{Der}(S)$  we then have  $x_2\partial_2 = D_{H,\omega}(x_2)^{[p]} \in \bar{\mathcal{G}}_0$ ; see [Str09, p. 45]. This means that  $\mathcal{G}_0 \cong \bar{\mathcal{G}}_0$  contains the degree derivation of  $\mathcal{G}$ . Since  $M_0$  is a restricted subalgebra of  $\mathfrak{g}$  we can find a toral element  $t_0 \in M_{(0)}$  which maps onto the degree derivation  $x_2\partial_2$  under the canonical homomorphism  $M_{(0)} \twoheadrightarrow \bar{\mathcal{G}}_0$ .

From [Str09, 10.4] we know that  $S_p = S \oplus \mathbb{k}(x_2\partial_2)$  and  $\bar{\mathcal{G}}_i = S_i$  for all  $i \neq 0$ . Since  $S_{-1} \neq 0$  there exists  $\varepsilon \in \{\pm 1\}$  such that  $S_i = \{x \in S \mid (g(1, t)) \cdot x = t^{\varepsilon i}x\}$  for all  $i \in \mathbb{Z}$ . As  $S$  is spanned by all  $D_{H,\omega}(f)$  with  $f \in O(1; \underline{1})$  by [Str04, Theorem 6.5.7(2)], this shows that  $\bar{\mathcal{G}}_i = S_i = 0$  for all  $i \leq -p$ . If  $\varepsilon = -1$  then  $S_i = 0$  for  $i > 1$ . So we can repeat verbatim the argument used in the last two paragraphs of (3.8) to obtain that  $N(\mathcal{G}) = 0$ . If  $\varepsilon = 1$  then  $S_i = 0$  for  $i > 1$  and we reach the same conclusion by arguing as in (3.13). In any event,  $\mathfrak{g}_{t_0} \subset M$  which implies that  $M$  is a regular subalgebra of  $\mathfrak{g}$ . Since this contradicts our assumption of  $M$  we deduce that the case  $S \cong H(2; \underline{1}; \Phi)^{(1)}$  is impossible.

3.24. Suppose  $S \cong H(2; (2, 1); \Phi)^{(2)}$ . Then  $\dim S \geq p^3 - 2$  which implies that  $G$  is of type  $E_7$  and  $p = 5$ . Therefore,

$$\dim N(\mathcal{G}) = \dim \mathfrak{g} - \dim \bar{\mathcal{G}} \leq \dim \mathfrak{g} - \dim S \leq 133 - 123 = 10.$$

From this it is immediate that  $S = S^{(\infty)}$  acts trivially on all  $\bar{\mathcal{G}}$ -modules  $N(\mathcal{G})^k/N(\mathcal{G})^{k+1}$  with  $k \geq 1$ . Let  $M'$  be the inverse image of  $S_0$  under the canonical homomorphism  $M = M_{(0)} \twoheadrightarrow \bar{\mathcal{G}}_0$ .

If  $\dim M' > (\dim \mathfrak{g})/2$  then  $\kappa|_{M'} \neq 0$ . Applying (9) then yields that the restriction of  $\kappa_{\mathcal{G}}$  to  $S_0$  is nonzero. Since  $S_0$  acts trivially on each  $N(\mathcal{G})^k/N(\mathcal{G})^{k+1}$ , it follows that the restriction of  $\kappa_{\bar{\mathcal{G}}}$  to  $S_0$  is nonzero as well. Since  $S$  is an ideal of  $\mathcal{G}$ , the restriction of  $\kappa_{\bar{\mathcal{G}}}$  to  $S$  coincides with the Killing form of  $S$ . In view of the above this means that  $\kappa_S \neq 0$ . But then  $\kappa_S$  is non-degenerate by the simplicity of  $S$ . Since  $S$  is strongly degenerate, i.e. contains a nonzero element  $c$  with  $(\text{ad } c)^2 = 0$ , this is impossible; see [PS01, Lemma 4.4], for example. It follows that

$$(18) \quad \dim S_{\geq 0} \leq \dim M' \leq (\dim \mathfrak{g})/2.$$

If  $S = H(2; (2, 1))^{(2)}$  we may assume without loss of generality that the grading of  $S$  has type  $(a_1, a_2)$  with respect to the standard generators  $x_1, x_2$  of  $O(2; (2, 1))$ ; see [PS01, Theorem 4.7]. Since  $S_{-1}$  is an irreducible  $\text{Der}_0(S)$ -module, the proof of [PS01, Lemma 4.18] shows that  $(a_1, a_2)$  is one of  $(-1, -1), (-1, 0), (0, -1), (0, 1), (1, 0), (1, 1)$ . The description in *loc. cit.* shows that in the last three cases  $S_{\geq 0}$  has codimension  $< p^2 = 25$  in  $S$ , implying that  $\dim S_{\geq 0} \geq 123 - 24 = 94 > (\dim \mathfrak{g})/2$ . In view of (18) it follows that  $(a_1, a_2) \in \{(-1, -1), (-1, 0), (0, -1)\}$ . Repeating almost verbatim the arguments used (3.18) and (3.19) we find, in each of the remaining cases, an element  $x \in M_{(1)} \cap \mathcal{O}_{\min}$  such that  $(\text{ad } \text{gr}_1(x))^4 \neq 0$ . Since  $(\text{ad } x)^3 = 0$  for every  $x \in \mathcal{O}_{\min}$ , we reach a contradiction. This shows that the case where  $S \cong H(2; (2, 1))^{(2)}$  cannot occur.

If  $S \cong H(2; (2, 1); \Phi(\tau))^{(1)}$  we argue as in (3.23). Indeed, in this case  $\text{Der}(S) = S_p$  by [Str04, Theorem 7.2.2(3)] and  $\omega = (1 + x_1^{(p^2-1)}x_2^{(p-1)})dx_1 \wedge dx_2$ . It follows from (17) that  $S_{-1}$  is an irreducible  $S_0$ -module. By the proof of Lemma 2.8, we may assume that the grading of  $S$  is induced by the action of the torus  $T_\Phi = \{g(t, t^{-p-1}) \mid t \in \mathbb{k}^\times\}$  whilst [Str04, Theorem 6.5.7(2)] yields that  $S$  is contained in the  $\mathbb{k}$ -span of all  $D_{H,\omega}(f)$  with  $f \in O(2; (2, 1))$ . Then  $S_0$  is contained in the  $\mathbb{k}$ -span of all

$D_{H,\omega}(x_1^{(kp+k-p)}x_2^{(k)})$  with  $k \in \{1, \dots, p-1\}$ . As a consequence, the nilradical of  $S_0$  has codimension 1 in  $S_0$ . As  $S_{-1}$  is an irreducible  $S_0$ -module  $S_{-1}$  this implies that  $\dim \mathcal{G}_{-1} = \dim S_{-1} = 1$ . But then  $\mathcal{G}_{-1}$  is not faithful over  $\mathcal{G}_0$ , a contradiction. Therefore,  $S \not\cong H(2; (2, 1); \Phi(\tau))^{(1)}$ .

If  $S \cong H(2; (2, 1); \Phi(1))$  then  $\omega = \exp(x_1^{(p^2)})dx_1 \wedge dx_2$  and we may assume that the grading of  $S$  is induced by the action of  $T_\Phi = \{g(1, t) \mid t \in \mathbb{k}^\times\}$  (see the proof of Lemma 2.8). Since  $S_{-1} \neq 0$  there exists  $\varepsilon \in \{\pm 1\}$  such that  $S_i = \{x \in S \mid (g(1, t)) \cdot x = t^{\varepsilon i} x\}$  for all  $i \in \mathbb{Z}$ . From [Str04, Theorem 7.2.(4)] we know that  $\text{Der}(S) = S_p$  whilst [Str04, Theorem 7.1.3(2)] implies that  $\text{Der}(S)$  is spanned by  $S \cong \text{ad } S$  and the elements  $x_2\partial_2$  and  $(\partial_1 - x_1^{(p^2-1)}x_2\partial_2)^p$  both of which have degree 0. By [Str04, Theorem 6.5.8],  $S$  is spanned by the homogeneous elements

$$D_{H,1}(f) = D_\omega(\exp(x_1^{(p^2)}) \cdot f), \quad f \in O(2; (2, 1)).$$

This implies that if  $\varepsilon = 1$  then the graded subalgebra  $S_{\geq 0}$  of  $\text{Der}(S)$  has codimension  $p^2 = 25$  in  $S$ . As  $\dim S_{\geq 0} = 100 > (\dim \mathfrak{g})/2$ , this violates (18). So it must be that  $\varepsilon = -1$ . But then the above description shows that  $S_0 \cong W(1; 2)$  as Lie algebras,  $\mathcal{G}_i = S_i = 0$  for  $i > 1$ , and  $\mathcal{G}_1 = S_1 = \text{span}\{D_{H,1}(x_1^{(i)}) \mid 0 \leq i \leq p^2 - 1\}$  is an irreducible (non-restrictable)  $S_0$ -module of dimension  $p^2$ . Using [Str04, (6.5.8)] it is straightforward to check that  $D_{H,1}(x_1^{(i)}) = x_1^{(i-1)}\partial_2$  for  $1 \leq i \leq p^2 - 1$  and  $D_{H,1}(1) = x_1^{(p^2-1)}\partial_2$ . Moreover, the element  $u := \partial_1 - x_1^{(p^2-1)}x_2\partial_2 \in S_0$  operates on  $S_1$  as a single Jordan block of size  $p^2$  with eigenvalue 1. Put  $v := \sum_{i=0}^{p^2-1} x_1^{(i)}\partial_2$ . Then  $v \in S_1$  and direct computations show that  $[u, v] = v$  and  $(\text{ad } v)^4 \neq 0$ . Since  $\mathcal{G}_i = 0$  for  $i > 1$ , the ideal  $\text{nil}(M) = M_{(1)}$  is abelian. Since  $\mathcal{G}_0$  acts faithfully on  $S_1$ , it coincides with  $\mathfrak{c}_{\mathfrak{g}}(\text{nil}(M))$ . Let  $\pi: M_{(0)} \rightarrow \mathcal{G}_0$  be the canonical homomorphism, and pick any  $\tilde{u} \in \pi^{-1}(u)$ . Let  $R$  be the Lie subalgebra of  $M$  generated by  $\tilde{u}$  and  $v$  (here we identify  $M_{(1)}$  with  $S_1$ ). Since  $[\tilde{u}, v] = v$  and  $v \in \text{nil}(M)$ , the above discussion in conjunction with Corollary 2.3 show that  $v \in \mathcal{O}_{\min}$ . But then  $(\text{ad } v)^3 = 0$ . As this contradicts our choice of  $v$  we conclude that  $S \not\cong H(2; (2, 1); \Phi(1))$ .

The case  $S \cong H(2; (2, 1); \Phi(2))$  is quite similar, but shorter. Here  $\omega = \exp(x_2^{(p)})dx_1 \wedge dx_2$  and the grading of  $S$  is induced by the action of  $T_\Phi = \{g(t, 1) \mid t \in \mathbb{k}^\times\}$ . Arguing as before we observe that  $\varepsilon = \pm 1$ . If  $\varepsilon = 1$  then  $S = \bigoplus_{i \geq -1} S_i$  and  $S_{\geq 0}$  has codimension  $p = 5$  in  $S$ . But then  $\dim S_{\geq 0} = 120 > (\dim \mathfrak{g})/2$  violating (18). If  $\varepsilon = -1$  then  $S_i = 0$  for  $i > 1$  and

$$S_{-1} = \text{span}\{D_\omega(\exp(x_2^{(p)}) \cdot x_1^{(2)}x_2^{(i)}) \mid 0 \leq i \leq p-1\}.$$

Using [Str04, (6.5.8)] it is straightforward to check that  $S_{-1} \subset W(2; \underline{1})$ . Since the Lie algebra  $S_{<0}$  is generated by  $S_{-1}$ , it follows that  $S_{1-p} \subset W(2; \underline{1})$  (here we regard both  $S$  and  $W(2; \underline{1})$  as Lie subalgebras of  $W(2; (2, 1))$ ). However,  $D_\omega(\exp(x_2^{(p)}) \cdot x_1^{(p)}) \in S_{1-p} \setminus W(2; \underline{1})$ . This contradiction finally shows that  $S$  is not a Lie algebra of type  $H$ .

3.25. If  $S = K(3; \underline{1})$  then  $S \subset W(3; \underline{1})$  has basis  $\{D_K(x_1^{a_1}x_2^{a_2}x_3^{a_3}) \mid 0 \leq a_i \leq p-1\}$ . In particular,  $\dim S = p^3$ . Hence this case may occur only if  $p = 5$  and  $G$  is a group of type  $E_7$ . Recall that if  $f \in O(3; \underline{1})$  then  $D_K(f) = f_1\partial_1 + f_2\partial_2 + f_3\partial_3$  where  $f_1 = x_1\partial_3(f) - \partial_2(f)$ ,  $f_2 = x_2\partial_3(f) + \partial_1(f)$  and  $f_3 = 2f - x_1\partial_1(f) - x_2\partial_2(f)$ ; see [BGP09, 2.11], for example.

Repeating verbatim the arguments used at the beginning of (3.24) we observe that  $\dim S_{\geq 0} \leq (\dim \mathfrak{g})/2$ . As all derivations of  $S$  are inner by [Str04, 7.1.2(4)], the  $S_0$ -module  $S_{-1}$  is faithful and irreducible. The Lie algebra  $S$  has basis  $\{D_K(f) \mid f \in O(3; \underline{1})\}$  and any  $\mathbb{Z}$ -grading of  $S$  is induced by an admissible grading of  $O(3; \underline{1})$ ; see [PS01, p. 276]. Hence it may be assumed without loss of



generality that there exists a triple  $(a_1, a_2, a_3) \in \mathbb{Z}^3$  with  $a_3 = a_1 + a_2$  such that

$$(19) \quad \deg(D_K(x_1^{r_1} x_2^{r_2} x_3^{r_3})) = (r_1 + r_3 - 1)a_1 + (r_2 + r_3 - 1)a_2.$$

We first suppose that  $a_2 = 0$ . Since  $S_{-1} \neq 0$  it follows from (19) that  $a_1 = \pm 1$  and  $S_0$  is spanned by  $D_K(x_1 x_2^i)$  and  $D_K(x_2^i x_3)$  with  $0 \leq i \leq p-1$ . Also,

$$\sum_{i < 0} S_{ia_1} = S_{-a_1} = \text{span}\{D_K(x_2^i) \mid 0 \leq i \leq p-1\}.$$

If  $a_1 = 1$  the  $\dim S_{\geq 0} = 120 > (\dim \mathfrak{g})/2$ . So this case cannot occur. Hence  $a_1 = -1$  implying  $S_{\geq 0} = S_0 \oplus S_1$ . Then the proof of [PS01, Lemma 4.12] shows that  $S_0 \cong W(1; \underline{1}) \ltimes O(1; \underline{1})$  acts irreducibly and faithfully of  $S_1 \cong O(1; \underline{1})$ . Set  $u := D_K(x_1)$  and  $v := D_K(x_2)$ . It is straightforward to see that  $u \in S_0$  and  $v$  spans  $S_{-1} \cap \ker(\text{ad } u)$ . Identify  $S_1$  with  $M_{(1)} = \text{nil}(M)$  and let  $\tilde{u}$  be a preimage of  $u$  in  $M = M_{(0)}$ . The preceding remarks show that  $\mathfrak{c}_{\mathfrak{g}}(\text{nil}(M)) = M_{(1)}$ . Applying Corollary 2.3 with the abelian Lie algebra  $R = \mathbb{k}\tilde{u} \oplus \mathbb{k}v$  we now conclude that  $v \in \mathcal{O}_{\min}$ . But  $(\text{ad } D_K(x_2))^4 \neq 0$  and  $(\text{ad } e)^3 = 0$  for all  $e \in \mathcal{O}_{\min}$ . This contradiction shows that  $a_1 \neq 0$ . Arguing similarly (with the roles of  $x_1$  and  $x_2$  interchanged) one obtains that  $a_2 \neq 0$ .

If  $a_1 \neq |a_2|$  then none of  $D_K(1)$ ,  $D_K(x_1)$ ,  $D_K(x_2)$ ,  $D_K(x_1^2)$ ,  $D_K(x_2^2)$  have degree 0. As  $D_K(x_3)$  and  $D_K(x_1 x_2)$  belong to the standard maximal subalgebra of  $S$  it follows that  $[S_0, S_0]$  acts nilpotently on  $S$ . The irreducibility of the  $\mathcal{G}_0$ -module  $\mathcal{G}_{-1} = S_{-1}$  then forces  $\dim \mathcal{G}_{-1} = 1$  yielding  $\mathcal{G}_k = 0$  for  $k \geq 2$ . But then  $\dim \mathfrak{g} = \dim S = 125$ , a contradiction. Therefore,  $a_1 = |a_2|$ .

If  $a_1 = -a_2$  then arguing as in part (d) of the proof of [PS01, Lemma 4.12] one observes that  $S_0$  contains a nonzero ideal acting nilpotently on  $S$ . As this contradicts the faithfulness of the  $S_0$ -module  $S_{-1}$ , we conclude that  $a_1 = a_2$ . Then  $a_1 = \pm 1$  because  $S_{-1} \neq 0$ . If  $a_1 = 1$  then  $\dim S_{\geq 0} = 122 > (\dim \mathfrak{g})/2$ . As this is not the case we have that  $a_1 = a_2 = -1$ . Then  $\mathcal{G}_2 = S_2 = \mathbb{k}D_K(1)$  and  $\mathcal{G}_k = S_k = 0$  for  $k \geq 2$ . It follows that  $M$  normalises a line  $\mathbb{k}e$  with  $e \in \mathcal{N}(\mathfrak{g})$ . As  $\mathfrak{n}_{\mathfrak{g}}(\mathbb{k}e)$  is contained in the optimal parabolic subalgebra of  $e$  by [Pre03, Theorem A], we now deduce that the case where  $S \cong K(3; \underline{1})$  is impossible.

3.26. Finally, suppose that  $S$  is isomorphic to the Melikian algebra  $\mathcal{M}(1, 1)$ . Then  $p = 5$  and  $\dim S = 125$  implying that the group  $G$  has type  $E_7$ . By [Str04, 7.1.4], all derivations of  $S$  are inner. As  $\dim \mathcal{G} = 125 < 133 = \dim \mathfrak{g}$  it must be that  $N(\mathcal{G}) \neq 0$ . Arguing as at the beginning of (3.24) we observe that  $\mathcal{G} = S = S^{(\infty)}$  acts trivially on the nonzero  $\mathcal{G}$ -module  $N(\mathcal{G})/N(\mathcal{G})^2$ . As a consequence,  $N(\mathcal{G})$  contains a graded ideal  $I$  of  $\mathcal{G}$  such that  $N(\mathcal{G})^2 \subset I$  and  $\dim(N(\mathcal{G})/I) = 1$ . Let  $L = \mathcal{G}/I$ . This Lie algebra is a central extension of  $\mathcal{M}(1, 1)$  and its centre  $\mathfrak{z}(L) = N(\mathcal{G})/I$  is contained in  $\mathcal{G}_{-k}/I_{-k}$  for some  $k \geq 2$ . Since  $\mathcal{G}_{-k} = \mathcal{G}_{-1}^k$  for all  $k \geq 1$  it must be that  $\mathfrak{z}(L) \subset [L, L]$ , i.e. the extension

$$0 \rightarrow \mathfrak{z}(L) \rightarrow L \rightarrow \mathcal{M}(1, 1) \rightarrow 0$$

is non-split. Since this contradicts [PS08, Proposition 6.2], we now conclude that the present case cannot occur. This completes the proof of Theorem 1.1.

#### 4. FURTHER REMARKS AND OBSERVATIONS

4.1. **Non-existence of Hamiltonian subalgebras of  $\mathfrak{g}$ .** We retain our assumption that  $G$  is an exceptional algebraic  $\mathbb{k}$ -group and  $p$  is a good prime for  $G$ . It is proved in [HS15b] that  $\mathfrak{g}$  does not contain Lie subalgebras  $M$  isomorphic to  $H(m; \underline{n}; \Phi)^{(2)}$ . Of course, the majority of pairs  $(m; \underline{n})$  are ruled out by simple dimension arguments, but the most difficult case where  $M \cong H(2; \underline{1}; \Phi)^{(2)}$  is addressed in *loc. cit.* by using a description of the Witt subalgebras of  $\mathfrak{g}$ . Since this result is important for classifying all maximal subalgebras of  $\mathfrak{g}$ , an alternative proof is given below.

**Proposition 4.1** (Herpel–Stewart). *The Lie algebra  $\mathfrak{g}$  does not contain Lie subalgebras isomorphic to  $H(2; \underline{1}; \Phi)^{(2)}$ .*

*Proof.* (a) Suppose  $S \cong H(2; \underline{1}; \Phi)^{(2)}$  is a Lie subalgebra of  $\mathfrak{g}$  and no exceptional groups  $H$  of dimension  $< \dim \mathfrak{g}$  has the property that  $H(2; \underline{1}; \Phi)^{(2)} \hookrightarrow \text{Lie}(H)$ . Let  $\mathcal{S}$  be the  $p$ -envelope of  $S$  in  $\mathfrak{g}$ . If  $\mathfrak{z}(\mathcal{S})$  contains a semisimple element,  $t$  say, then  $S$  is contained in the Levi subalgebra  $\mathfrak{g}_t = \text{Lie}(L)$  where  $L = G_t$ . If  $\mathfrak{z}(\mathcal{S})$  contains a nonzero nilpotent element,  $e$  say, then  $M \subset \mathfrak{g}_e$ . Since  $\mathfrak{g}_e$  is contained in a proper parabolic subalgebra of  $\mathfrak{g}$ , the Lie algebra  $S$  again injects into the Lie algebra of a proper Levi subgroup  $L$  of  $G$ . In any event, the Lie algebra of one of the simple components of  $L$  must contain an isomorphic copy of  $S$ . As this component must be classical by our assumption on  $\mathfrak{g}$ , it is straightforward to see that  $S$  affords a faithful representation of dimension  $< 23 \leq p^2 - 2$ . However, our discussion in (3.20) shows that this is impossible. Therefore,  $\mathfrak{z}(\mathcal{S}) = 0$  that is  $\mathcal{S} \cong S_p$ , the minimal  $p$ -envelope of  $S$ ; see [Str04, Corollary 1.1.8(2)]. In particular, this implies that all composition factors of the  $S$ -module  $\mathfrak{g}$  are restrictable.

(b) Since all composition factors of the  $\mathcal{S}$ -module  $\mathfrak{g}$  are restricted, the arguments used in (3.17) show that  $\mathcal{S}$  contains a nonzero  $p$ -balanced toral element, say  $h$ . Then the centraliser  $\mathfrak{g}_h$  is listed in one of the six cases of Proposition 2.7. Let  $S = S_{(-1)} \supset S_{(0)} \supset S_{(1)} \supset \cdots \supset S_{(q)}$  be the standard filtration of the Cartan type Lie algebra  $S$  and denote by  $O(2; \underline{1})[d]$  the subspace of all homogeneous truncated polynomials of degree  $d$  in  $O(2; \underline{1})$ . It is well known that  $q \in \{2p-5, 2p-4\}$  and  $\dim S_{(i)}/S_{(i+1)} \geq \dim O(2; \underline{1})[i+2]$  for all  $i \geq -1$ . From this it is immediate that  $S_{(p-1)}$  is an abelian subalgebra of  $\mathfrak{g}$  and

$$\dim S_{(p-1)} \geq \sum_{i \geq p+1} O(2; \underline{1})[i] = (p-2) + (p-3) + \cdots + 2 = \frac{1}{2}(p-1)(p-2) - 1.$$

If  $p = 11$  then  $\dim S_{(p-1)} \geq 44$  and if  $p = 7$  then  $\dim S_{(p-1)} \geq 14$ . Applying [Ma45] and [PeS15] and arguing as in (3.21) we now observe that cases (v) and (vi) of Proposition 2.7 cannot occur.

(c) Suppose  $S \cong H(2; \underline{1})^{(2)}$ . Some dimension estimates used in (3.20) and (3.21) are applicable in the present case since they only rely on the structure of  $S$  and properties of its irreducible restricted representations. So we adopt the notation introduced in (3.20). In particular, we let  $E_1, \dots, E_r$  stand for the *induced* composition factors of the  $\text{ad}_{\mathfrak{g}} S$ -module  $\mathfrak{g}$  and write  $l$  for the multiplicity of the adjoint module  $S$  in  $\mathfrak{g}$ . As in (3.20) we put

$$s := l + \sum_{i=1}^r (\dim E_i)/p^2.$$

Recall that  $S$  is closely related with  $O(2; \underline{1})$  regarded as a Lie algebra through its standard Poisson bracket  $\{\cdot, \cdot\}$ . More precisely,  $\mathbb{k}1 \subset \{O(2; \underline{1}), O(2; \underline{1})\}$  is the centre of the Poisson algebra  $O(2; \underline{1})$  and  $S \cong \{O(2; \underline{1}), O(2; \underline{1})\}/\mathbb{k}1$  as Lie algebras. The derived subalgebra  $\{O(2; \underline{1}), O(2; \underline{1})\}$  is spanned by all monomials  $x_1^{m_1} x_2^{m_2}$  with  $0 \leq m_1, m_2 \leq p-1$  and  $m_1 + m_2 < 2(p-1)$ . In other words, we may identify  $S$  with the  $\mathbb{k}$ -span of all  $x_1^{m_1} x_2^{m_2} \in O(2; \underline{1})$ , with  $0 < m_1 + m_2 < 2(p-1)$  in such a way that

$$(20) \quad [x_1^{m_1} x_2^{m_2}, x_1^{n_1} x_2^{n_2}] = (m_1 n_2 - m_2 n_1) x_1^{m_1+n_1-1} x_2^{m_2+n_2-1} \quad (0 \leq m_i, n_i \leq p-1, i = 1, 2).$$

Let  $h = -2(1 + x_1)x_2$ , a toral element of  $S$  not contained in  $S_{(0)}$ . Using (20) it is easy to see that all nonzero eigenvalues on  $\text{ad } h$  have multiplicity  $p$ . If  $\rho: S \rightarrow \mathfrak{gl}(E)$  is a restricted representation of  $S$  induced from  $S_{(0)}$ , so that  $E = u(S) \otimes_{u(S_{(0)})} E_0$  for some irreducible restricted  $S_{(0)}$ -module  $E_0$ , then *all* eigenvalues of  $\rho(h)$  have multiplicity  $p \dim E_0$  because  $h \notin S_{(0)}$  and  $h^{[p]} = h$ . From this it is immediate that  $h$  is a  $p$ -balanced element of  $\mathfrak{g}$  and  $\dim \mathfrak{g}(h, i) = ps$  for all  $i \in \mathbb{F}_p^\times$ .

(d) Let  $V$  be the 2-dimensional subspace of  $S$  spanned by  $x_1$  and  $(1 + x_1)^2 x_2$ . By construction,  $V \cap S_{(0)} = 0$  and  $[h, v] = 2v$  for all  $v \in V$ . Analysing Kac coordinates of  $p$ -balanced  $G$ -orbits obtained in the course of proving Proposition 2.7 one finds out that in cases (i), (ii) and (iv) there always exist a nilpotent element  $e \in \mathcal{O}(A_{p-1})$  and an optimal cocharacter  $\tau: \mathbb{k}^\times \rightarrow G$  for  $e$  such that  $e \in \mathfrak{g}(\tau, 2)$  and  $h = (d\tau)(1)$ . In order to see this one just needs to compare (3), (4) and (7) with [Car93, pp. 402–407]. Moreover, in all three case  $\mathfrak{g}(h, 2) = \mathfrak{g}((d\tau)(1), 2) = \mathfrak{g}(\tau, 2) \oplus \mathbb{k}n$  for some  $n \in \mathcal{O}_{\min}$ . As  $\dim V = 2$  it must be that  $V \cap \mathfrak{g}(\tau, 2) \neq 0$ . In view of [Pre03, Theorem 2.3] this implies that a nonzero element  $v \in V$  is contained in the Zariski closure of  $\mathcal{O}(A_{p-1})$  in  $\mathfrak{g}$ . Note that  $v = \lambda x_1 + \mu(1 + x_1)^2 x_2 = v = u + v_0$  where  $v_0 = 2\mu x_1 x_2 + x_1 x_2^2 \in S_{(0)}$  and  $u = \lambda x_1 + \mu x_2$  is a nonzero element in the linear span of  $x_1 = \partial_2$  and  $x_2 = -\partial_1$ .

As  $\mathcal{O}(A_{p-1}) \subset \mathcal{N}_p(\mathfrak{g})$  we have that  $v^{[p]} = 0$ . Since  $v \notin S_{(0)}$  it is straightforward to see that for any irreducible restricted representation  $\rho: S \rightarrow \mathfrak{gl}(E)$  such that  $E = u(S) \otimes_{u(S_{(0)})} E_0$  as  $S$ -modules, the endomorphism  $\rho(v)$  has  $p \dim E_0$  Jordan blocks of size  $p$ . On the other hand, using (20) one checks directly that the  $\text{Im}(\text{ad}_S u)^{p-1}$  has dimension  $p - 1$ . A standard filtration argument then shows that  $\dim \text{Im}(\text{ad}_S v)^{p-1} \geq \dim \text{Im}(\text{ad}_S u)^{p-1}$ . Since  $(\text{ad } v)^p = 0$  and  $\dim S = p^2 - 2$  it follows that  $\text{ad}_S v$  has  $p - 1$  Jordan blocks of size  $p$ . Consequently,  $\text{ad } v$  has at least  $s(p - 1)$  Jordan blocks of size  $p$ .

Regarding  $v = \lambda x_1 + \mu(1 + x_1)^2 x_2$  as an element of the Poisson algebra  $O(2; \underline{1})$  we observe that  $v^k \in \{O(2; \underline{1}), O(2; \underline{1})\}$  for all  $1 \leq k \leq p - 1$ . We identify the  $v, v^2, \dots, v^{p-1}$  with their images in  $S \cong \{O(2; \underline{1}), O(2; \underline{1})\}/\mathbb{k}1$ . Since  $\{v, v^k\} = 0$  and  $[h, v] = 2v$  we have that  $[v, v^k] = 0$  and  $[h, v^k] = 2kv^k$  for all  $1 \leq k \leq p - 1$ . This implies that  $\mathfrak{g}(h, i) \cap \mathfrak{g}_v \neq 0$  for all  $i \in \mathbb{F}_p^\times$ .

Note that  $v^{p-1}, u^{p-1} \in S_{(p-3)}$  and  $v^{p-1} - u^{p-1} \in S_{(p-2)}$ . Since the automorphism group  $\text{Aut}(S)$  preserves all components  $S_{(i)}$  of the standard filtration of  $S$ , it is straightforward to see that  $u^{p-1}$  lies in the Zariski closure of  $\mathbb{k}^\times (\text{Aut}(S) \cdot v^{p-1})$ . In view of Chevalley's Semicontinuity Theorem, this means that  $\dim \text{Im}(\text{ad}_S v^{p-1}) \geq \dim \text{Im}(\text{ad}_S u^{p-1})$ . Since for every irreducible restricted representation  $\rho: S \rightarrow \mathfrak{gl}(E)$  induced from  $S_{(0)}$  the vector space  $E$  carries an  $\text{Aut}(S)$ -module structure compatible with that of  $S$ , we also have, by the same token, that  $\dim \text{Im} \rho(v^{p-1}) \geq \dim \text{Im} \rho(u^{p-1})$ . Since  $u^{p-1}$  and  $x_1^{p-1}$  are conjugate under the action of  $\text{Aut}(S)$  the images of  $\text{ad}_S(u^{p-1})$  and  $\text{ad}_S(x_1^{p-1})$  have equal dimensions. The same applies to the images of  $\rho(u^{p-1})$  and  $\rho(x_1^{p-1})$  thanks to the compatible action of  $\text{Aut}(S)$  on  $E$ .

(e) If  $h$  is as in case (i) of Proposition 2.7 then  $p = 5$ , the group  $G$  has type  $E_6$  and  $s = 3$ . So  $\text{ad } v$  has at least 12 Jordan block of size  $p$ . If  $h$  is as in case (ii) of Proposition 2.7 then  $p = 5$ , the group  $G$  has type  $E_7$  and  $s = 5$ . Therefore,  $\text{ad } v$  has at least 20 Jordan blocks of size  $p$ . If  $h$  is as in case (iii) of Proposition 2.7 then  $p = 7$ , the group  $G$  has type  $E_8$  and  $s = 5$ . Therefore,  $\text{ad } v$  has at least 30 Jordan blocks of size  $p$ . Since  $(\text{Ad } G)v \subseteq \overline{\mathcal{O}(A_{p-1})}$  (and hence  $\dim [\mathfrak{g}, v] \leq \dim \mathcal{O}(A_{p-1})$ ) we combine [Law95, Tables 6, 8 and 9] with [PSt15, Theorem 4.1] to conclude that either  $v \in \mathcal{O}(A_{p-1})$  or  $G$  is of type  $E_8$  and  $v \in \mathcal{O}(E_8(a_7))$ .

Suppose  $v \in \mathcal{O}(E_8(a_7))$ . Then  $p = 7$  and  $v$  is distinguished in  $\mathfrak{g}$ . So it follows from [Bor91, 11.8] that all maximal toral subalgebras of the normaliser  $\mathfrak{n}_{\mathfrak{g}}(\mathbb{k}v) = \text{Lie}(N_G(\mathbb{k}v))$  are 1-dimensional and conjugate under the adjoint action of  $G_v$ . By [LT11, p. 157], there is a nonzero toral element  $t \in \mathfrak{n}_{\mathfrak{g}}(\mathbb{k}v)$  such that  $\mathfrak{g}(t, 2p - 2) \cap \mathfrak{g}_v = 0$ . As  $[h, v] = 2v$  the above implies that  $\mathfrak{g}(h, i) \cap \mathfrak{g}_v = 0$  for some  $i \in \mathbb{F}_p^\times$ , a contradiction.

Now suppose  $v \in \mathcal{O}(A_{p-1})$ . Then it is immediate from [LT11, pp. 87, 104, 158] that the subspace  $\mathfrak{g}_v \cap \mathfrak{g}(h, 2p-2) = \mathfrak{g}_v \cap \mathfrak{g}((d\tau)(1), 2p-2) = \mathfrak{g}_v \cap \mathfrak{g}(\tau, 2p-2)$  is contained in the Zariski closure of  $\mathcal{O}_{\min}$ . Our earlier remarks then show that  $v^{p-1} \in \mathcal{O}_{\min}$ . In particular,  $\dim \operatorname{Im}(\operatorname{ad} v^{p-1}) < 2(p-1)s$  in all three cases. On the other hand, arguing as in (3.20) and taking into account the last paragraph of part (d) one observes that  $\dim \operatorname{Im} \operatorname{ad}(v^{p-1}) \geq 2(p-1)s$ . This contradiction shows that cases (i), (ii) and (iv) of Proposition 2.7 cannot occur.

If  $h$  is as in case (iii) of Proposition 2.7, then it follows from (6) that  $\mathfrak{g}(h, i) = 0$  for some  $i \in \mathbb{F}_p^\times$ . However,  $0 \neq v^k \in \mathfrak{g}(h, 2k)$  for all  $k \in \{1, \dots, p-1\}$ . This proves that  $S \not\cong H(2; \underline{1})^{(2)}$ .

(f) Suppose  $S \cong H(2; \underline{1}; \Phi(\tau))^{(1)}$ . The rule  $\{x_1, x_2\}_\omega = (1+x_1)(1+x_2)$  extends uniquely to a Poisson bracket on  $O(2; \underline{1})$  and  $S$  can be identified with  $(O(2; \underline{1})/\mathbb{k}1, \{\cdot, \cdot\}_\omega)$  as Lie algebras. Moreover,  $S_p = \operatorname{Der}(S) = T \oplus S$  where  $T = \mathbb{k}(1+x_1)\partial_1 \oplus \mathbb{k}(1+x_2)\partial_2$ ; see [Str09, p. 42] (as before, we identify  $S$  with  $\operatorname{ad} S$ ). We have shown in part (a) that  $S_p$  is isomorphic to the  $p$ -envelope of  $S$  in  $\mathfrak{g}$ . It is well known that decomposing  $S$  into weight spaces with respect to  $T$  enables one to identify  $S$  with the Block algebra  $\operatorname{Bl}(\mathbb{F}_p \oplus \mathbb{F}_p)$  which is the  $\mathbb{k}$ -span of the symbols  $v_\alpha$  for  $\alpha \in (\mathbb{F}_p \oplus \mathbb{F}_p) \setminus \{0\}$  with Lie bracket given by

$$(21) \quad [v_\alpha, v_\beta] = (\alpha|\beta)v_{\alpha+\beta}$$

where  $(\cdot|\cdot)$  is a non-degenerate skew-symmetric  $\mathbb{F}_p$ -valued bilinear form on  $\mathbb{F}_p \oplus \mathbb{F}_p$ ; see [Str09, Theorem 10.3.2]. Here  $\mathbb{F}_p \oplus \mathbb{F}_p$  is nothing but the dual space of  $T^{\operatorname{tor}} = \mathbb{F}_p(1+x_1)\partial_1 \oplus \mathbb{F}_p(1+x_2)\partial_2$  and  $(\mathbb{F}_p \oplus \mathbb{F}_p) \setminus \{0\}$  identifies with the set of all weights of  $T$  on  $S$ . By [BW82, Lemma 4.6.4] or by the proof of [Str09, Theorem 10.7.3], any non-trivial irreducible restricted  $S_p$ -module  $E$  has  $p^2 - 1$  nonzero  $T$ -weights, all of the same multiplicity (depending on  $E$ ). From this it is immediate that any nonzero toral element of  $T$  is  $p$ -balanced in  $\mathfrak{g}$ .

Pick linearly independent  $\alpha, \beta \in \mathbb{F}_p \oplus \mathbb{F}_p$  and set  $V = \mathbb{k}v_\alpha \oplus \mathbb{k}v_\beta$ , a 2-dimensional subspace of  $\mathfrak{g}$ . There is a toral element  $h \in T$  with  $\alpha(h) = \beta(h) = 2$ . This element is  $p$ -balanced in  $\mathfrak{g}$  by the preceding remark. By [Str09, Theorem 10.3.2(5)], there exist linearly independent toral elements  $t_\alpha, t_\beta \in T$  such that  $v_\alpha^{[p]} = at_\alpha$  and  $v_\beta^{[p]} = bt_\beta$  for some  $a, b \in \mathbb{k}^\times$ . Combining (21) with Jacobson's formula for  $p$ -th powers one observes that

$$(\lambda v_\alpha + \mu v_\beta)^{[p]} - \lambda^p at_\alpha - \mu^p bt_\beta \in \sum_{\gamma \in (\mathbb{F}_p \oplus \mathbb{F}_p) \setminus \{0\}} \mathbb{k}v_\gamma$$

for all  $\lambda, \mu \in \mathbb{k}$ . This implies that  $V \cap \mathcal{N}_p(\mathfrak{g}) = 0$ . On the other hand, it follows from (3), (4), (6) and (7) that  $\mathfrak{g}(h, 2) \cap \mathcal{N}_p(\mathfrak{g})$  contains a subspace of codimension  $\leq 1$  in  $\mathfrak{g}(h, 2)$ . Since  $V \subset \mathfrak{g}(h, 2)$  is 2-dimensional we reach a contradiction thereby proving that the case  $S \cong H(2; \underline{1}; \Phi(\tau))^{(1)}$  is impossible.

(g) Finally, suppose  $S \cong H(2; \underline{1}; \Phi(1))$ . It is well known that  $S$  is isomorphic to an Albert–Zassenhaus algebra  $L(\Gamma, \Theta)$  where  $\Gamma$  is an additive subgroup of  $\mathbb{k}$  and  $\Theta: \Gamma \rightarrow \mathbb{k}$  is a group homomorphism. Recall that  $L(\Gamma, \Theta)$  is spanned by the symbols  $u_\alpha$  with  $\alpha \in \Gamma$  and

$$(22) \quad [u_\alpha, u_\beta] = (\beta - \alpha + \alpha\Theta(\beta) - \beta\Theta(\alpha))u_{\alpha+\beta}.$$

All isomorphism types of Albert–Zassenhaus algebras of a given dimension are determined in [BIO79]. Since in the present case  $\dim S = |\Gamma| = p^2$ , we may assume further that  $\Gamma = \mathbb{F}_{p^2}$ , the set of all roots of  $X^{p^2} - X = 0$  in  $\mathbb{k}$ , and  $\Theta = \operatorname{Fr}$ , the Frobenius automorphism of  $\mathbb{F}_{p^2}$ ; see [BIO79, Corollary 5.3]. The  $[p]$ -powers  $u_\alpha^{[p]} = (\operatorname{ad} u_\alpha)^p$  with  $\alpha \in \mathbb{F}_{p^2}$  span a self-centralising 2-dimensional torus  $T$  of  $S_p \cong \operatorname{Der}(S)$  such that  $T \cap S = \mathbb{k}u_0$ . It has  $p^2$  weights on  $S$  each of multiplicity 1; see [Str09, Theorem 10.4.6]. The corresponding weight spaces are nothing but  $\mathbb{k}u_\beta$  with

$\beta \in \mathbb{F}_{p^2}$ . Since  $[u_0, u_\beta] = \beta u_\beta$  for all  $\beta \in \mathbb{F}_{p^2}$  by (22), it is straightforward to see that  $u_0^{[p]} \notin \mathbb{k}u_0$ . Hence  $T = \mathbb{k}u_0 \oplus \mathbb{k}u_0^{[p]}$  and  $T \cap S$  is not a restricted subalgebra of  $S_p$ . Applying [BW82, Lemma 4.8.1] we now deduce that any non-trivial irreducible restricted  $S_p$ -module  $E$  has  $p^2 - 1$  nonzero  $T$ -weights, all of the same multiplicity (depending on  $E$ ). As before, this implies that any nonzero toral element of  $T$  is  $p$ -balanced in  $\mathfrak{g}$ . It follows from (22) that

$$(23) \quad \begin{aligned} [u_\alpha, u_{\beta+k\alpha}] &= (\beta + (k-1)\alpha + \alpha(\beta^p + k\alpha^p) - (\beta + k\alpha)\alpha^p)u_{\beta+(k+1)\alpha} \\ &= ((\beta(1 - \alpha^p) + \alpha(\beta^p + k - 1))u_{\beta+(k+1)\alpha}. \end{aligned}$$

for all  $\alpha, \beta \in \mathbb{F}_{p^2}$  and all  $k \in \mathbb{F}_p$ . Taking  $\alpha = 1$ , the identity element of  $\mathbb{F}_{p^2}$ , we get

$$[u_1, u_{\beta+k}] = (\beta^p + k - 1)u_{\beta+k+1} \quad (\forall \beta \in \mathbb{F}_{p^2}).$$

Therefore,  $[u_1^{[p]}, u_\beta] = \prod_{i \in \mathbb{F}_p} (\beta^p + i) \cdot u_\beta = (\beta^{p^2} - \beta^p)u_\beta = (\beta - \beta^p)u_\beta$  for all  $\beta \in \mathbb{F}_{p^2}$  which yields  $u_1^{[p]} = u_0 - u_0^{[p]}$ . This means that  $u_1^{[p]} \neq 0$  and  $(\text{ad } u_1)^{p-1}(u_\beta) \neq 0$  whenever  $\beta \notin \mathbb{F}_p$ . Applying (23) with  $(\alpha, \beta) = (\beta, 1)$  we get

$$[u_\beta, u_{1+k\beta}] = (1 + k\beta - \beta^p)u_{1+(k+1)\beta} \quad (\forall \beta \in \mathbb{F}_{p^2}).$$

Note that  $\Theta^2 = \text{Fr}^2 = \text{Id}$ . If  $\beta \in \mathbb{F}_{p^2}$  is such that  $\beta^p = -\beta$  then  $[u_\beta, u_{1+k\beta}] = (1 + (k+1)\beta)u_{1+(k+1)\beta}$ , implying

$$\begin{aligned} (\text{ad } u_\beta)^{p-1}(u_1) &= \left(\prod_{i \in \mathbb{F}_p^\times} (1 + i\beta)\right)u_{1-\beta} = \beta^{p-1} \left(\prod_{i \in \mathbb{F}_p^\times} (\beta^{-1} + i)\right)u_{1-\beta} \\ &= \beta^{p-1}(\beta^{1-p} - 1)u_{1-\beta} = (1 - \beta^{p-1})u_{1-\beta} = \beta^{-1}(\beta - \beta^p)u_{1-\beta} = 2u_{1-\beta}. \end{aligned}$$

It follows that  $[u_\beta^{[p]}, u_1] = 2[u_\beta, u_{1-\beta}] = 2u_1$ . In particular,  $u_\beta^{[p]} \neq 0$ .

Let  $V$  be the  $\mathbb{k}$ -span of  $u_1$  and  $u_\beta$  where  $\beta^p = -\beta$ . As  $[u_0, u_1] = [u_0^{[p]}, u_1] = u_1$  and  $[u_0, u_\beta] = -[u_0^{[p]}, u_\beta] = \beta u_\beta$ , the  $T$ -weights of  $u_1$  and  $u_\beta$  are  $\mathbb{F}_p$ -independent. Therefore, they form a basis of  $T^*$ . As a consequence, there exists a nonzero toral element  $h \in T$  such that  $[h, v] = 2v$  for all  $v \in V$ . This element must be  $p$ -balanced in  $\mathfrak{g}$  by our earlier remarks. Jacobson's formula for  $p$ -th powers shows that

$$(\lambda u_1 + \mu u_\beta)^{[p]} - \lambda^{p-1}\mu(\text{ad } u_1)^{p-1}(u_\beta) - \lambda\mu^{p-1}(\text{ad } u_\beta)^{p-1}(u_1) \in T \oplus \sum_{\alpha \notin \{0, \pm(\beta-1)\}} \mathbb{k}u_\alpha$$

for all  $\lambda, \mu \in \mathbb{k}$ . Since both  $(\text{ad } u_1)^{p-1}(u_\beta) \in S_{\beta-1}$  and  $(\text{ad } u_\beta)^{p-1}(u_1) \in S_{1-\beta}$  are nonzero by our choice of  $\beta$  we now deduce that  $V \cap \mathcal{N}_p(\mathfrak{g}) = 0$  (one should also keep in mind here that  $u_1^{[p]} \neq 0$  and  $u_\beta^{[p]} \neq 0$ ). At this point we can argue as at the end of in part (f) to conclude that  $S \not\cong H(2; \underline{1}; \Phi(1))$ . This completes the proof.  $\square$

**4.2. Counterexamples to Morozov's theorem in bad characteristic: type  $E_8$ .** Suppose  $G$  is a group of type  $E_8$  and  $p = 5$ . Although 5 is a bad prime for  $G$  there is a bijection between the nilpotent orbits in  $\mathfrak{g} = \text{Lie}(G)$  and  $\mathfrak{g}_{\mathbb{C}} = \text{Lie}(G_{\mathbb{C}})$  which preserves the orbit dimensions. In the notation of [Pre03, 2.6], each  $G$ -orbit  $\mathcal{O} \subset \mathcal{N}(\mathfrak{g})$  has the form  $\mathcal{O} = \mathcal{O}(I, J)$  for a suitable pair  $(I, J) \in \mathcal{P}(\Pi)$  and by [CP13, Theorem 1.4] each cocharacter  $\lambda_{I,J} \in X_*(G)$  constructed in [Pre03, 2.6] is still optimal in the sense of the Kempf–Rousseau theory for a nice representative  $e_{I,J} \in \mathcal{O}(I, J)$ . In particular, the stabiliser  $G_{e_{I,J}}$  is contained in the parabolic subgroup  $P(\lambda_{I,J})$ .

Due to the above-mentioned bijection between the nilpotent orbits of  $\mathfrak{g}$  and  $\mathfrak{g}_{\mathbb{C}}$  each Hesselink stratum of  $\mathcal{N}(\mathfrak{g})$  (i.e. each Lusztig's nilpotent piece of  $\mathfrak{g}$ ) is a single  $G$ -orbit; see [CP13] for detail. This means that for any  $(I, J) \in \mathcal{P}(\Pi)$  the orbit  $(\text{Ad } P(\lambda_{I,J})) \cdot e_{I,J}$  is Zariski dense in  $\bigoplus_{i \geq 2} \mathfrak{g}(\lambda_{I,J}, i)$



and  $\text{Lie}(G_{e_{I,J}}) \subseteq \text{Lie}(P(\lambda_{I,J}))$ . Comparing the tables in [VAG05, pp. 90–93] and [Car93, pp. 405–407] one observes that the elements  $e \in \mathcal{N}(\mathfrak{g})$  for which  $\text{Lie}(G_e) \subsetneq \mathfrak{g}_e$  lie in two orbits which have Dynkin labels  $E_8$  and  $A_4 + A_3$ . That  $\mathcal{O}(E_8)$  has this property was first pointed out in [Spr66, Theorem 5.9]. Note that  $\mathcal{O}(A_4 + A_3) \subset \mathcal{N}_p(\mathfrak{g})$  has dimension 200 whilst [VAG05, p. 92] implies that  $\dim \mathfrak{g}_e = 50$  for any  $e \in \mathcal{O}(A_4 + A_3)$ . It turns out that in characteristic 5 the centraliser  $\mathfrak{g}_e$  has very unusual properties.

By [VAG05, p. 86], the standard representative  $e = \sum_{\alpha \in \Pi \setminus \{\alpha_5\}} e_\alpha$  still lies in  $\mathcal{O}(A_4 + A_3)$  which in conjunction with [CP13, Theorem 5.2] shows that the cocharacter  $\tau \in X_*(G)$  from [LT11, p. 148] is still optimal for  $e$ . Using [LT07, pp. 237, 238] it is not hard to check that  $\mathfrak{g}_e(\tau, 0)$  is spanned by an  $\mathfrak{sl}_2$ -triple  $\{e_0, h_0, f_0\}$  and  $\mathfrak{g}_e(\tau, 1)$  is an irreducible module for  $\mathfrak{g}_e(\tau, 0) \cong \mathfrak{sl}_2$  of highest weight 3. In [LMT09, §4] (which also deals with  $\mathcal{O}(A_4 + A_3)$  in characteristic 5) one finds two nonzero elements  $X, Y \in \mathfrak{g}_e(\tau, -1)$  and checks directly that  $[h_0, X] = X$  and  $[h_0, Y] = -Y$ . On the other hand, the preceding remarks entail that  $\text{Lie}(G_e) \subseteq \bigoplus_{i \geq 0} \mathfrak{g}_e(\tau, i)$  has dimension 48. As  $\dim \mathfrak{g}_e = 50$  we now deduce that

$$\mathfrak{g}_e = \bigoplus_{i \geq -1} \mathfrak{g}_e(\tau, i), \quad \text{Lie}(G_e) = \bigoplus_{i \geq 0} \mathfrak{g}_e(\tau, i), \quad \dim \mathfrak{g}_e(\tau, -1) = 2.$$

Since the vectors from  $\bigoplus_{i \geq 0} \mathfrak{g}_e(\tau, i)$  displayed in [LT07, pp. 237, 238] are still linearly independent in characteristic 5, they form a basis for  $\text{Lie}(G_e)$ .

Set  $h := (d\tau)(1)$  and  $\mathfrak{n}_e := \mathfrak{n}_{\mathfrak{g}}(\mathbb{k}e)$ . As  $[h, e] = 2e$  we have that  $\mathfrak{n}_e = \mathbb{k}h \oplus \mathfrak{g}_e$ . As the torus  $\tau(\mathbb{k}^\times)$  acts on  $\mathfrak{g}_e$  by Lie algebra automorphisms the radical of  $\mathfrak{g}_e$  is a graded subspace of  $\mathfrak{g}_e = \bigoplus_{i \geq -1} \mathfrak{g}_e(\tau, i)$ . In particular, it is  $(\text{ad } h)$ -stable. At the author's request Thomas Purslow has checked the following by using some standard GAP routines:

- (i) the radical  $A$  of  $\mathfrak{g}_e$  is abelian and has dimension 24;
- (ii) the Lie subalgebra  $\mathfrak{g}'_e$  of  $\mathfrak{g}_e$  generated by  $\mathfrak{g}(\tau, 1)$  and  $\mathfrak{g}(\tau, -1)$  has dimension 47;
- (iii) the normaliser  $\mathfrak{w} := \mathfrak{n}_{\mathfrak{g}}(A)$  has dimension 74;
- (iv) the Lie algebra  $\mathfrak{w}/A$  is simple and restricted.

**Theorem 4.2.** *The following are true for any  $e \in \mathcal{O}(A_4 + A_3)$ :*

- (1)  $A \subset \mathfrak{g}'_e$  and  $\mathfrak{g}'_e/A \cong H(2; \underline{1})^{(2)}$  as Lie algebras.
- (2)  $A = \text{rad}(\mathfrak{g}_e)$  and  $\mathfrak{g}_e/A \cong H(2; \underline{1})$  as Lie algebras.
- (3)  $A = \text{rad}(\mathfrak{n}_e)$  and  $\mathfrak{n}_e/A \cong \text{Der}(H(2; \underline{1})^{(2)})$  as Lie algebras.
- (4)  $A = \text{rad}(\mathfrak{w})$  and  $\mathfrak{w}/A \cong W(2; \underline{1})$  as Lie algebras.
- (5)  $A \cong (O(2; \underline{1})/\mathbb{k}1)^*$  as  $W(2; \underline{1})$ -modules.
- (6)  $A \subset \mathcal{N}(\mathfrak{g})$  and  $\mathfrak{w}$  is a maximal Lie subalgebra of  $\mathfrak{g}$ .

*Proof.* (a) Using [LT07, p. 237] one observes that  $\mathfrak{g}_e(\tau, 1)$  is an irreducible 4-dimensional  $\mathfrak{g}_e(\tau, 0)$ -module. If  $[\mathfrak{g}_e(\tau, -1), \mathfrak{g}_e(\tau, 1)] = 0$  then the Engel–Jacobson theorem implies that  $\sum_{i \neq 0} \mathfrak{g}_e(\tau, i)$  is a nilpotent ideal of codimension 3 in  $\mathfrak{g}_e$ . As this contradicts (i) it must be that  $[\mathfrak{g}_e(\tau, -1), \mathfrak{g}_e(\tau, 1)] = \mathfrak{g}_e(\tau, 0)$  and  $A \subseteq \bigoplus_{i \geq 2} \mathfrak{g}_e(\tau, i)$ . Let  $L = \mathfrak{g}_e/A$ . As  $A$  is a graded subspace of  $\mathfrak{g}_e$  we have that  $L = \bigoplus_{i \geq -1} L_i$  where  $L_i = \mathfrak{g}_e(\tau, i)/A \cap \mathfrak{g}_e(\tau, i)$ . Also,  $L_0 \cong \mathfrak{sl}_2$ , and  $L_{-1}$  is an irreducible  $L_0$ -module. Since  $L$  is semisimple it satisfies the conditions of the Weak Recognition Theorem; see [BGP09, Theorem 2.66]. As  $L$  is a restricted Lie algebra,  $L_0 \cong \mathfrak{sl}_2$  and  $L_1 \not\cong L_{-1}^*$  applying that theorem shows that  $L$  is sandwiched between  $H(2; \underline{1})^{(2)}$  and  $H(2; \underline{1})$ . As a consequence, the Lie subalgebra

$L'$  generated by  $L_{\pm 1}$  is isomorphic to  $H(2; \underline{1})^{(2)}$ , so that  $\dim L' = p^2 - 2 = 23$ . In conjunction with (i) and (ii) this gives  $A \subset \mathfrak{g}'_e$  proving (1).

(b) As  $L \subseteq H(2; \underline{1})$  and  $\dim H(2; \underline{1}) = p^2 + 1 = 26 = \dim \mathfrak{g}_e - \dim A = \dim L$  by [BGP09, 2.10] and (i), the equality must hold, i.e.  $L = H(2; \underline{1})$ . This proves (2). Since  $\mathfrak{n}_e = \mathbb{k}h \oplus \mathfrak{g}_e$  and  $A$  is  $(\text{ad } h)$ -stable, we have that  $A = \text{rad}(\mathfrak{n}_e)$ . In this situation [BGP09, Theorem 2.66] applies to the graded Lie algebra  $\mathfrak{n}_e/A$  forcing

$$H(2; \underline{1}) \cong L \subsetneq \mathfrak{n}_e/A \subseteq CH(2; \underline{1}).$$

As  $CH(2; \underline{1}) \cong \text{Der}(H(2; \underline{1})^{(2)})$  by [Str09, Theorem 7.1.2(2)], statement (3) follows.

(c) As  $L_{-i} = 0$  for  $i > 1$  it also follows from [BGP09, Theorem 2.66] that the grading of the Cartan type Lie algebra  $L = H(2; \underline{1})$  is standard. In particular,  $L_k = 0$  for  $k \geq 2p - 4 = 6$ . In view of [LT07, p. 237] this means that  $A$  contains  $\mathfrak{g}_e(\tau, 9)$ , a 2-dimensional irreducible  $\mathfrak{g}_e(\tau, 0)$ -module contained in the centre of the Lie algebra  $\bigoplus_{i \geq 0} \mathfrak{g}_e(\tau, i)$ . Since  $\mathbb{k}e$  is a trivial submodule of the  $L$ -module  $A$  and  $A/\mathbb{k}e$  has dimension  $23 = p^2 - 2$ , we can apply [RH98, Corollary 510] to conclude that  $A/\mathbb{k}e \cong L'$  as  $L$ -modules.

Since no Lie algebras of the form  $\mathfrak{psl}_{kp}$  or  $\text{Lie}(\mathcal{H})$ , where  $\mathcal{H}$  is a simple algebraic  $\mathbb{k}$ -group, have dimension  $50 = \dim \mathfrak{w}/A$ , the Classification Theorem from [PS08] implies that  $\mathfrak{w}/A$  is isomorphic to a restricted Lie algebra of Cartan type. As  $\dim W(m; \underline{1}) = m5^m$  for  $m \geq 1$ ,  $\dim S(m; \underline{1})^{(1)} = (m-1)(5^m - 1)$  for  $m \geq 3$ ,  $\dim H(2m; \underline{1})^{(2)} = 5^{2m} - 2$  for  $m \geq 1$ , and  $\dim K(2m+1, \underline{1}) \geq 5^{2m+1} - 1$  for  $m \geq 1$ , we have only one option here, namely,  $\mathfrak{w}/A \cong W(2; \underline{1})$ . This proves (4).

(d) Since  $A$  is a graded subspace of  $\mathfrak{g}_e$  and  $A \cap \mathfrak{g}_e(\tau, i) = 0$  for  $i = \{-1, 0, 1\}$  by our remarks in part (a) we have that  $A \subseteq \bigoplus_{i \geq 2} \mathfrak{g}_e(\tau, i)$ . In particular,  $A \subset N(\mathfrak{g})$ . Since  $\mathfrak{w}$  is  $(\text{Ad } \tau(\mathbb{k}^\times))$ -stable and  $A(\tau, i) = 0$  for  $i \leq 1$ , it must be that  $[\mathfrak{w}(\tau, k), e] = 0$  for  $k \leq -1$ . This forces  $\mathfrak{w} = \mathfrak{g}_e(\tau, -1) \oplus (\bigoplus_{i \geq 0} \mathfrak{w}(\tau, i))$ . As  $W(2; \underline{1})$  has a unique subalgebra of codimension 2 by Kreknin's theorem we now obtain that the image of  $\bigoplus_{i \geq 0} \mathfrak{w}(\tau, i)$  in  $\mathfrak{w}/A$  coincides with the standard maximal subalgebra of  $\mathfrak{w}/A \cong W(2; \underline{1})$ . Consequently,  $\mathfrak{w}(\tau, 0) = \mathfrak{n}_e(\tau, 0) \cong \mathfrak{gl}_2$ .

Since  $A$  is a non-trivial restricted  $W(2; \underline{1})$ -module of dimension  $p^2 - 1$ , it follows from [RH01, Theorem 4.2] that either  $A \cong O(2; \underline{1})/\mathbb{k}1$  or  $A \cong (O(2; \underline{1})/\mathbb{k}1)^*$  as  $W(2; \underline{1})$ -modules. In the former case the annihilator  $A^{\mathfrak{g}_e(\tau, -1)}$  of  $\mathfrak{g}_e(\tau, -1) \cong \mathbb{k}\partial_1 \oplus \mathbb{k}\partial_2$  in  $A$  is a 2-dimensional irreducible module over  $\mathfrak{g}_e(\tau, 0) \cong \mathfrak{sl}_2$  which generates an irreducible  $L$ -submodule isomorphic to  $L'$ . Since  $\mathbb{k}e$  is a trivial  $L$ -submodule of  $A$ , this would imply that  $\dim A^{\mathfrak{g}_e(\tau, -1)} \geq 3$  contrary to the fact that  $(O(2; \underline{1})/\mathbb{k}1)^{\partial_1} \cap (O(2; \underline{1})/\mathbb{k}1)^{\partial_2}$  is 2-dimensional. So  $A \cong (O(2; \underline{1})/\mathbb{k})^*$  proving (5).

(e) It remains to show that  $\mathfrak{w}$  is a maximal subalgebra of  $\mathfrak{g}$ . Suppose this is not the case and let  $\mathcal{L}$  be a proper restricted Lie subalgebra of  $\mathfrak{g}$  with  $\mathfrak{w} \subsetneq \mathcal{L}$ . If  $I$  is a nonzero ideal of  $\mathcal{L}$  then  $I^A \neq 0$  and hence  $A \subseteq I$  by the irreducibility of the  $\mathfrak{w}$ -module  $A$ . If  $\text{nil}(\mathcal{L}) \neq 0$  then we can take  $\mathfrak{z}(\text{nil}(\mathcal{L})) \neq 0$  for  $I$  to conclude that  $\text{nil}(\mathcal{L})$  is a nilpotent ideal of  $\mathfrak{g}_e$ . But then  $\text{nil}(\mathcal{L}) = A$ , by part (a), forcing  $\mathcal{L} \subseteq \mathfrak{w}$ . As this contradicts our choice of  $\mathcal{L}$  we deduce that  $\text{nil}(\mathcal{L}) = 0$ .

Suppose  $\text{rad}(\mathcal{L}) \neq 0$ . Then  $\mathcal{L}$  contains a nonzero abelian restricted ideal. Since  $\text{nil}(\mathcal{L}) = 0$  this ideal must contain a nonzero  $[p]$ -semisimple element of  $\mathfrak{g}$ , say  $z$ . As  $(\text{ad}_{\mathcal{L}} z)^2 = 0$  and  $\text{Ker } \text{ad}_{\mathcal{L}} z = \text{Ker } (\text{ad}_{\mathcal{L}} z)^2$  by the semisimplicity of  $\text{ad}_{\mathcal{L}} z$ , it must be that  $\mathfrak{z}(\mathcal{L}) \neq 0$ . Taking  $I = \mathfrak{z}(\mathcal{L})$  we then obtain that  $A \subseteq \mathfrak{z}(\mathcal{L})$  which, in turn, yields  $\mathcal{L} \subseteq \mathfrak{g}_e$ , a contradiction.

Thus from now on we may assume that  $\mathcal{L}$  is semisimple. If  $\mathcal{L}$  contains two distinct minimal ideals  $I_1$  and  $I_2$  then the above shows that  $A \subseteq I_1 \cap I_2 = 0$ , a contradiction. Therefore,  $\mathcal{L}$  contains a unique

minimal ideal,  $J$  say. In this situation Block's theorem says that there exist a simple Lie algebra  $S$  and a nonnegative integer  $m$  such that  $J \cong S \otimes O(m; \underline{1})$  as Lie algebras and

$$S \otimes O(m; \underline{1}) \subset \mathcal{L} \subseteq (\operatorname{Der}(S) \otimes O(m; \underline{1})) \rtimes (\operatorname{Id}_S \otimes W(m; \underline{1})).$$

Furthermore, since  $\mathcal{L}$  is restricted the image of  $\mathcal{L}$  under the canonical projection

$$\pi: (\operatorname{Der}(S) \otimes O(m; \underline{1})) \rtimes (\operatorname{Id}_S \otimes W(m; \underline{1})) \rightarrow W(m; \underline{1})$$

is a transitive subalgebra of  $W(m; \underline{1})$ , i.e. does not preserve any nonzero proper ideals of the commutative algebra  $O(m; \underline{1})$ ; see [Str04, Corollary 3.3.5]. Since  $\dim \mathfrak{g} = 248$ ,  $\dim S \geq 3$  and  $\dim O(n; \underline{m}) = 5^m$  we have that  $m \in \{0, 1, 2\}$ .

(f) Suppose  $m > 0$  and let  $\mathfrak{m}$  be the maximal ideal of the local algebra  $O(m; \underline{1})$ . If  $\pi(\mathfrak{w}) = 0$  then  $\mathfrak{w} \subseteq \operatorname{Der}(S) \otimes O(m; \underline{1})$ . In this case the abelian ideal  $J_0 := S \otimes \mathfrak{m}^{m(p-1)}$  of  $J$  is an  $(\operatorname{ad} \mathfrak{w})$ -stable and  $J_0^A$  is a nonzero  $\mathfrak{w}$ -submodule of  $J_0$ . But then  $J_0^A \subset \mathfrak{g}_e \subset \mathfrak{w}$  is an abelian ideal of  $\mathfrak{w}$  implying  $J_0^A = A$ . As  $J_0^A \subseteq \mathfrak{z}(S \otimes \mathfrak{m})$  and  $S \otimes \mathfrak{m} = \operatorname{nil}(J)$  this entails that  $S \otimes \mathfrak{m}$  is a nilpotent ideal of  $\mathfrak{g}_e$  containing  $A$ . But then  $S \otimes \mathfrak{m} = A = J_0^A \subseteq S \otimes \mathfrak{m}^{m(p-1)}$ , a contradiction. We thus deduce that  $\pi(\mathfrak{w})$  is a nonzero Lie subalgebra of  $W(m; \underline{1})$ .

Since  $A \subset J$  by part (e) and  $W(2; \underline{1}) \cong \mathfrak{w}/A$  is a simple Lie algebra, the above discussion shows that  $\pi$  identifies  $W(2; \underline{1})$  with a nonzero Lie subalgebra of  $W(m; \underline{1})$ . The concluding remark in part (e) now shows that  $m = 2$  and the restriction of  $\pi$  to  $\mathfrak{w}$  is surjective. Since in this situation  $\dim S < 10$  it is immediate from [PS08, Theorem 1.1] that  $\operatorname{Der}(S) = \operatorname{ad} S \cong S$ , so that

$$\mathcal{L} = J + \mathfrak{w} \cong (S \otimes O(2; \underline{1})) \rtimes (\operatorname{Id}_S \otimes (W(2; \underline{1}))) \quad \text{and} \quad J \cap \mathfrak{w} = A.$$

The algebra  $O(2; \underline{1})$  is spanned by the monomials  $x^{\mathbf{a}} := x_1^{a_1} x_2^{a_2}$  with  $0 \leq a_i \leq p-1$ . As  $e \in S \otimes O(2; \underline{1})$  we can write  $e = s_0 \otimes 1 + \sum_{|\mathbf{a}| \geq 1} s_{\mathbf{a}} \otimes x^{\mathbf{a}}$  for some  $s_0, s_{\mathbf{a}} \in S$ . As  $A$  is an irreducible  $\mathfrak{w}$ -module and  $W(2; \underline{1})$  has no Lie subalgebras of codimension 1, it follows from Kreknin's theorem that  $\mathfrak{w}$  contains a unique Lie subalgebra of codimension 2, namely, the inverse image of the standard maximal subalgebra  $W(2; \underline{1})_{(0)}$  under the canonical homomorphism  $\mathfrak{w} \twoheadrightarrow W(2; \underline{1})$ ; we call it  $\mathfrak{w}_{(0)}$ . It is immediate from our discussion in part (d) that  $\mathfrak{w} = \mathfrak{g}_e(\tau, -1) \oplus \mathfrak{w}_{(0)}$ . This yields that  $\pi(\mathfrak{g}_e(\tau, -1))$  contains elements of the form  $\partial_1 + u_1$  and  $\partial_2 + u_2$  for some  $u_1, u_2 \in W(2; \underline{1})_{(0)}$ . As  $e$  commutes with  $\mathfrak{g}_e(\tau, -1)$  this implies that  $s_0 \neq 0$ .

Let  $\mathbf{p} - \mathbf{1} = (p-1, p-1)$ . Since  $s_0 \otimes x^{\mathbf{p}-\mathbf{1}}$  commutes with  $e$  we have that  $s_0 \otimes x^{\mathbf{p}-\mathbf{1}} \in \mathfrak{g}_e \cap J \subseteq \mathfrak{w} \cap J = A$ . Applying to  $s_0 \otimes x^{\mathbf{p}-\mathbf{1}} \neq 0$  the endomorphisms  $\operatorname{ad} y$  with  $y \in \mathfrak{g}_e(\tau, -1)$  we observe that for any  $\mathbf{a} = (a_1, a_2)$  with  $0 \leq a_1, a_2 \leq p-1$  the  $\mathfrak{w}$ -module  $A$  contains an element of the form  $s_0 \otimes x^{\mathbf{a}} + \sum_{|\mathbf{b}| > |\mathbf{a}|} t_{\mathbf{b}} \otimes x^{\mathbf{b}}$  for some  $t_{\mathbf{b}} \in S$ . But then  $p^2 - 1 = \dim A \geq \dim O(2; \underline{1}) = p^2$ . This contradiction shows that  $m = 0$ .

(g) From now on we may assume that any proper restricted Lie subalgebra  $\mathcal{L}$  of  $\mathfrak{g}$  containing  $\mathfrak{w}$  is sandwiched between  $S \cong \operatorname{ad} S$  and  $\operatorname{Der}(S)$  for some simple Lie algebra  $S$ . It follows from [Str04, Chapter 7] and the Classification Theorem proved in [PS08] that the Lie algebra of outer derivations of  $S$  is solvable. Since the Lie algebra  $\mathfrak{w}$  is perfect we have that  $\mathfrak{w} \subset S$ . Therefore, no generality will be lost by assuming that  $\mathfrak{w}$  is a maximal subalgebra of  $S$ .

Suppose  $S \cong \mathfrak{psl}_{5k}$  for some  $k \geq 1$ . Since  $\dim S > 74$  by (iii) and  $5k - 2 = \operatorname{TR}(S) \leq \operatorname{TR}(\mathcal{L}) \leq \operatorname{TR}(\mathfrak{g}) = 8$  by [Pre87], [Pre90] and [Str04, Theorem 1.2.7] it must be that  $k = 2$ . Since  $\operatorname{Der}(S) \cong \mathfrak{pgl}_{10}$  by [BGP09, Lemma 2.7] and  $\operatorname{TR}(\mathfrak{pgl}_{10}) = 9$  by [Pre87] and [Pre90] the restricted Lie algebras  $\mathcal{L}$  and  $\mathfrak{psl}_{10}$  are isomorphic (one should keep in mind here that  $\mathcal{L}$  contains  $\operatorname{ad} S$  which has codimension 1 in  $\operatorname{Der}(S)$ ). As a consequence,  $\mathfrak{psl}_{10}$  contains a maximal toral subalgebra of  $\mathfrak{g}$ , say  $\mathfrak{t}$ . In view

of [Hum67, Theorem 13.3] or [Bor91, 11.8] there is a maximal torus  $T$  in  $G$  with  $\mathfrak{t} = \text{Lie}(T)$ . Since  $p > 3$ , it follows from [Sel67, Ch. II, § 3] that all root spaces of  $\mathfrak{g}$  with respect to  $\mathfrak{t}$  are 1-dimensional and hence  $(\text{Ad } T)$ -stable. In conjunction with [Sel67, Ch. II, § 4] this implies the set  $\Phi_0 := \{\gamma \in \Phi(G, T) \mid \mathcal{L}_\gamma \neq 0\}$  is a root subsystem of the root system  $\Phi = \Phi(G, T)$ . As  $|\Phi_0| = 90$ , it is immediate from the Borel–de Siebenthal theorem that  $\Phi_0$  is contained in a maximal root subsystem of type  $E_7 + A_1$  or  $D_8$  in  $\Phi$ . The simplicity of the restricted Lie algebra  $\mathcal{L}$  then entails that it embeds into a restricted Lie algebra  $\mathfrak{g}_0$  of type  $E_7$  or  $D_8$ . The first possibility cannot occur as Lie algebras of type  $E_7$  do not contain 8-dimensional toral subalgebras. Hence  $\mathfrak{g}_0 \cong \mathfrak{so}(V)$  where  $V$  is a 16-dimensional vector space over  $\mathbb{k}$ . Let  $\rho: \mathcal{L} \rightarrow \mathfrak{gl}(V)$  be the representation of  $\mathcal{L}$  induced by inclusion  $\mathcal{L} \subset \mathfrak{so}(V)$ . As  $\dim \mathcal{L} > 74 > 60 = (\dim \mathfrak{g}_0)/2$ , the trace form  $(X, Y) \mapsto \text{trace}((\rho(X) \circ \rho(Y)))$  on  $\mathcal{L}$  is nonzero and hence non-degenerate by the simplicity of  $\mathcal{L} \cong \mathfrak{psl}_{10}$ . Since this contradicts [Bl62] we now conclude that  $S \not\cong \mathfrak{psl}_{5k}$ .

Suppose  $S \cong \text{Lie}(\mathcal{H})$  for some simple algebraic  $\mathbb{k}$ -group  $\mathcal{H}$ . Since  $\dim S < 248$  and  $S$  is simple, 5 is a very good prime for  $\mathcal{H}$ . Theorem 1.1 then says that  $\mathfrak{w} = \text{Lie}(\mathcal{P})$  for some maximal parabolic subgroup  $\mathcal{P}$  of  $\mathcal{H}$ . In view of our discussion at the beginning of part (e) this implies that  $A = \text{Lie}(R_u(\mathcal{P}))$ . Then  $\text{Lie}(\mathcal{P})/\text{Lie}(R_u(\mathcal{P})) \cong W(2; \underline{1})$  as Lie algebras. Since  $\text{Lie}(\mathcal{P})/\text{Lie}(R_u(\mathcal{P}))$  is isomorphic to the Lie algebra of a reductive  $\mathbb{k}$ -group this is false.

Therefore,  $S$  is a Lie algebra of Cartan type (not necessarily restricted). Since  $\mathfrak{g} \cong \mathfrak{g}^*$  as  $(\text{Ad } G)$ -modules, the Lie algebra  $\mathfrak{g}$  admits a non-degenerate  $\mathfrak{g}$ -invariant bilinear form,  $b$  say. We stress that  $b \neq \kappa$  as in the present case the Killing form of  $\mathfrak{g}$  is identically zero. The explicit formulae in [CP13, p. 661] show that the form  $b$  is symmetric. If  $\dim S > 124$  then the restriction of  $b$  to  $S$  is nonzero and hence non-degenerate by the simplicity of  $S$ . Hence  $\mathfrak{g} = S \oplus E$  where  $E = \{x \in \mathfrak{g} \mid b(x, S) = 0\}$ . Since  $b$  is  $S$ -invariant,  $[S, E] \subseteq E$ . Since  $\mathfrak{w} \subset S$  by our earlier remarks and  $E^A \neq 0$  is  $(\text{ad } \mathfrak{w})$ -stable, the irreducibility of the  $\mathfrak{w}$ -module  $A$  yields  $A \subseteq E$ . But then  $A \subseteq S \cap E = 0$ , a contradiction. So  $74 < \dim S \leq 124$ . The results proved in [Str04, §§ 6.3–6.7] now imply that  $S \cong H(2; (2, 1), \Phi)^{(2)}$  where  $\Phi \in \{\text{Id}, \Phi(\tau)\}$ . In any event, the standard maximal subalgebra  $S_{(0)}$  of  $S$  is a restricted Lie subalgebra of  $\text{Der}(S)$  and  $S_{(0)}/\text{nil}(S_{(0)}) \cong \mathfrak{sl}_2$ ; see [Str04, Theorems 7.2.2 and 6.3.10].

Since  $\mathfrak{w} \subset S$  and  $S_{(0)}$  has codimension 2 in  $S$ , the Lie algebra  $\mathfrak{w} \cap S_{(0)}$  has codimension  $\leq 2$  in  $\mathfrak{w}$ . Since  $S_{(0)}$  has no homomorphic images isomorphic to  $W(2; \underline{1})$  and  $W(2; \underline{1})$  has no subalgebras of codimension 1, our discussion in part (f) shows that  $\mathfrak{w} \cap S_{(0)} = \mathfrak{w}_{(0)}$ . Note that  $\text{nil}(\mathfrak{w}_{(0)})$  coincides with the preimage of  $W(2; \underline{1})_{(1)}$  under the canonical homomorphism  $\mathfrak{w}_{(0)} \twoheadrightarrow W(2; \underline{1})_{(0)}$  and  $\mathfrak{w}_{(0)}/\text{nil}(\mathfrak{w}_{(0)}) \cong \mathfrak{gl}_2$ . By our remark at the beginning of this part, the  $p$ -closure  $\mathcal{S}$  of  $S$  in  $\mathfrak{g}$  is semisimple. Therefore,  $\mathcal{S} \cong S_p$  as restricted Lie algebras, where  $S_p$  denotes the  $p$ -envelope of  $S$  in  $\text{Der}(S)$ . It follows that  $\mathfrak{w}_{(0)}$  is a *restricted* subalgebra of  $S_{(0)}$ . But then the restricted subalgebra

$$(\mathfrak{w}_{(0)} + \text{nil}(S_{(0)})/\text{nil}(S_{(0)}) \cong \mathfrak{w}_{(0)}/(\mathfrak{w}_{(0)} \cap \text{nil}(S_{(0)}))$$

of  $S_{(0)}/\text{nil}(S_{(0)}) \cong \mathfrak{sl}_2$  has a copy of  $\mathfrak{gl}_2$  as a homomorphic image. As this is obviously false, we conclude that  $S$  does not exist. Thus  $\mathfrak{w}$  is a maximal subalgebra of  $\mathfrak{g}$  and our proof is complete.  $\square$

**Conjecture 4.3.** *Suppose  $G$  is of type  $E_8$  and  $p = 5$ . We conjecture that any maximal subalgebra  $M$  of  $\mathfrak{g}$  with  $\text{rad}(M) \neq 0$  is either conjugate to  $\mathfrak{w}$  under the adjoint action of  $G$  or has the form  $M = \text{Lie}(P)$  for some maximal parabolic subgroup  $P$  of  $G$  or coincides with the centraliser of a toral element  $t$  of  $\mathfrak{g}$ . In the latter case,  $M = \text{Lie}(G_t)$  and  $G_t$  is a semisimple group of type  $A_4A_4$ .*

*Remark 4.4.* It is quite possible that the maximal subalgebra  $\mathfrak{w}$  gives rise to a Weisfeiler filtration  $\mathcal{F}$  of  $\mathfrak{g}$  such that the corresponding graded Lie algebra  $\text{gr}_{\mathcal{F}}(\mathfrak{g})$  is isomorphic to the special Cartan type Lie algebra  $S(3; \underline{1})^{(1)}$ . This guess is supported by numerology:  $\dim S(3; \underline{1})^{(1)} = 2(p^3 - 1) = 248$ , and

the fact that the Lie algebra  $\mathcal{G} = S(3; \underline{1})^{(1)}$  does admit a  $\mathbb{Z}$ -grading  $\mathcal{G} = \bigoplus_{i \in \mathbb{Z}} \mathcal{G}_i$  with  $\mathcal{G}_i = 0$  for  $i \geq 2$ ,  $\mathcal{G}_0 \cong W(2; \underline{1})$  and  $\mathcal{G}_1 \cong (O(2; \underline{1})/\mathbb{k}1)^*$  as  $W(2; \underline{1})$ -modules; see [PS01, p. 283]. This grading did not feature in our proof of Theorem 1.1 because the possibility that  $S \cong S(3; \underline{1})^{(1)}$  was ruled out in 3.15 by dimension arguments.

**4.3. Non-existence of Melikian subalgebras of  $\mathfrak{g}$ .** We continue assuming that  $G$  is a group of type  $E_8$  and  $p = 5$ . In this subsection we are going to show that no Melikian algebras  $M = \mathcal{M}(m, n)$  can occur as subalgebras of  $\mathfrak{g}$ . Since  $\dim \mathcal{M}(m, n) = 5^{m+n+1}$  and  $\dim \mathfrak{g} = 248$ , the Lie algebras  $\mathcal{M}(m, n)$  with  $m + n \geq 3$  are immediately ruled out by dimension reasons.

Suppose  $M = \mathcal{M}(1, 1)$  is a Lie subalgebra of  $\mathfrak{g}$  and let  $\mathcal{M}$  be the  $p$ -closure of  $M$  in  $\mathfrak{g}$ . The centre  $\mathfrak{z}(\mathcal{M})$  is an abelian restricted subalgebra of  $\mathfrak{g}$  and hence decomposes as  $\mathfrak{z}(\mathcal{M}) = \mathfrak{z}(\mathcal{M})_s \oplus \mathfrak{z}(\mathcal{M})_n$  where  $\mathfrak{z}(\mathcal{M})_s$  and  $\mathfrak{z}(\mathcal{M})_n$  are the maximal torus and the nilradical of  $\mathfrak{z}(\mathcal{M})$ , respectively. If  $\mathfrak{z}(\mathcal{M})_s \neq 0$  then  $\mathcal{M}$  is contained in the centraliser of a nonzero toral element of  $\mathfrak{g}$ , say  $t$ . Since  $t \in \text{Lie}(T)$  for some maximal torus  $T$  of  $G$ , by [Bor91, 11.8], and  $M$  is a simple Lie algebra of dimension 125, it is immediate from the Borel–de Siebenthal theorem that  $M$  embeds into a Lie algebra of type  $E_7$ . If  $\mathfrak{z}(\mathcal{M})_n \neq 0$  then  $\mathcal{M}$  is contained in the centraliser of a nonzero element  $e \in \mathcal{N}_p(\mathfrak{g})$ . Since  $\dim \mathfrak{g}_e \geq 125$  we have that  $\mathcal{O}(e) \neq \mathcal{O}(A_4 + A_3)$ . Then our discussion at the beginning of (4.2) shows that  $\mathfrak{g}_n$  is contained in a proper parabolic subalgebra  $\mathfrak{p} = \text{Lie}(P)$  of  $\mathfrak{g}$ . Since  $P = L \cdot R_u(P)$  for some Levi subgroup  $L$  and  $\text{Lie}(R_u(P))$  is nilpotent, the simple Lie algebra  $M$  injects into  $\mathfrak{l} := \text{Lie}(L)$ . Since  $\dim L \geq 125$  the Lie algebra  $[\mathfrak{l}, \mathfrak{l}]$  must have type  $E_7$ .

A Lie algebra  $\mathfrak{g}_0$  of type  $E_7$  admits a non-degenerate invariant symmetric bilinear form  $\eta: \mathfrak{g}_0 \times \mathfrak{g}_0 \rightarrow \mathbb{k}$ . If  $M \subset \mathfrak{g}_0$  the restriction of  $\eta$  to  $M$  is nonzero because  $\dim M > (\dim \mathfrak{g}_0)/2$ . The simplicity of  $M$  then shows that  $\eta|_M$  is a non-degenerate. But then  $\mathfrak{g}_0 = M \oplus M^\perp$  where  $M^\perp$  is the orthogonal complement of  $M$  in  $\mathfrak{g}_0$ . As  $\dim M^\perp = \dim \mathfrak{g}_0 - 125 = 133 - 125 = 8$  we have that  $\dim M > 64 = \dim \mathfrak{gl}(M^\perp)$ . As a consequence, the simple Lie algebra  $M$  must act trivially on  $M^\perp$ , that is  $[M, M^\perp] = 0$ . But then  $[\mathfrak{g}_0, M] = [M + M^\perp, M] \subseteq M$  which implies that  $M$  is an ideal of  $\mathfrak{g}_0$ . This contradiction shows that  $\mathfrak{z}(\mathcal{M}) = 0$ . Since  $M$  is a restricted Lie algebra, applying [Str04, Corollary 1.1.8] yields  $\mathcal{M} \cong M_p \cong M$ . Therefore,  $\mathcal{M} = M$ , i.e.  $M$  is a restricted subalgebra of  $\mathfrak{g}$ .

Let  $b$  be the invariant symmetric bilinear form on  $\mathfrak{g}$  introduced at the end of the proof of Theorem 4.2. Since  $M$  is simple and  $\dim M > (\dim \mathfrak{g})/2$  the restriction of  $b$  to  $M$  is non-degenerate. Therefore,  $\mathfrak{g} = M \oplus M^\perp$  where  $M^\perp = \{x \in \mathfrak{g} \mid b(x, M) = 0\}$ . As  $[M, M^\perp] \subseteq M^\perp$ , the preceding discussion shows that  $M^\perp$  is a restricted  $M$ -module. If all composition factors of the  $M$ -module  $M^\perp$  are trivial then  $[M, M^\perp] = [M^{(\infty)}, M^\perp] = 0$  implying that  $M$  is an ideal of  $\mathfrak{g}$ . As this contradicts the simplicity of  $\mathfrak{g}$  we may assume that at least one composition factor of the  $M$ -module  $M^\perp$  is non-trivial. In particular,  $M^\perp$  is a faithful  $M$ -module.

It was first observed by Kuznetsov in [Ku91] that the restricted Melikian algebra  $M$  admits a  $(\mathbb{Z}/3\mathbb{Z})$ -grading  $M = M_{\bar{0}} \oplus M_{\bar{1}} \oplus M_{\bar{2}}$  such that  $M_{\bar{0}}$  is a restricted Lie subalgebra of  $M$  isomorphic to  $W(2; \underline{1})$ . Given a 2-dimensional torus  $\mathfrak{t}_0$  of  $M_{\bar{0}}$  we can decompose the restricted  $M_{\bar{0}}$ -module  $M^\perp$  into weight spaces with respect to  $\mathfrak{t}$ . We denote by  $\Gamma(M^\perp, \mathfrak{t}_0)$  the set of all *nonzero*  $\mathfrak{t}$ -weights of  $M^\perp$ . Since the Lie algebra  $M_{\bar{0}}$  is simple and  $M^\perp$  is a faithful  $M$ -module, at least one composition factor of the  $M_{\bar{0}}$ -module  $M^\perp$  is non-trivial. Applying [Str09, 10.7.3] we now deduce that  $\Gamma(\mathfrak{t}_0, M^\perp) = p^2 - 1 = 24$ . By [Pre94, Lemma 4.1], there exists a 2-dimensional torus  $\mathfrak{t}_0$  in  $M_{\bar{0}}$  whose centraliser  $H$  in  $M$  is a 5-dimensional Cartan subalgebra of  $M$  with the property that  $[H, [H, H]] = \mathfrak{t}_0$ . Each weight space  $M_\gamma^\perp$  with  $\gamma \in \Gamma(\mathfrak{t}_0, M^\perp)$  is invariant under the adjoint action of  $H$ . Since  $\gamma \in \mathfrak{t}_0^* \setminus \{0\}$



and  $\mathfrak{t}_0 \subseteq [H, H]$ , the  $H$ -module  $M_\gamma^\perp$  has no 1-dimensional composition factors. Representation theory of nilpotent Lie algebras then yields that each  $\mathfrak{t}_0$ -weight space of  $M^\perp$  has dimension divisible by  $p = 5$ . As a consequence,  $\dim M^\perp \geq p(p^2 - 1) = 120$ . Since  $\dim M^\perp = 240 - 125 = 123$  and  $\dim \mathfrak{c}_M(\mathfrak{t}_0) = \dim H = 5$  this gives

$$\dim \mathfrak{c}_\mathfrak{g}(\mathfrak{t}_0) = \dim \mathfrak{c}_M(\mathfrak{t}_0) + \dim \mathfrak{c}_{M^\perp}(\mathfrak{t}_0) \leq 5 + 3 = 8.$$

Since  $\mathfrak{t}_0$  is a toral subalgebra of  $\mathfrak{g}$  and the field  $\mathbb{k}$  is infinite we can find an element  $t \in \mathfrak{t}_0$  such that  $\mathfrak{c}_\mathfrak{g}(\mathfrak{t}_0) = \mathfrak{g}_t$ . By [Bor91, 11.8], there exists a maximal torus  $T$  of  $G$  such that  $t \in \text{Lie}(T)$ . Since  $\text{Lie}(T)$  is abelian and has dimension 8 it must be that  $\mathfrak{c}_\mathfrak{g}(\mathfrak{t}_0) = \mathfrak{g}_t = \text{Lie}(T)$ . Since  $H \subset \mathfrak{c}_\mathfrak{g}(\mathfrak{t}_0)$  is non-abelian, this is impossible. We have reached a contradiction thereby proving that  $\mathcal{M}(1, 1)$  is not isomorphic to a Lie subalgebra of  $\mathfrak{g}$ .

**4.4. Counterexamples to Morozov's theorem in bad characteristic: type  $G_2$ .** In this subsection we assume that  $G$  is a group of type  $G_2$  and  $p = 2$ . It was first observed by Robert Steinberg in [St61, 2.6] that  $\mathfrak{g} \cong \mathfrak{psl}_4$ . This curious fact is explained in [CE15, Sect. 4] as follows: in characteristic 2, the irreducible  $G$ -module  $E = L(\varpi_1)$  of highest weight  $\varpi_1$  has dimension 6 and carries a nonzero  $G$ -invariant symplectic form. This gives us an inclusion  $\mathfrak{g} \subseteq \mathfrak{sp}(E)$ . As the Lie algebra  $\mathfrak{g}$  is simple we then have  $\mathfrak{g} \subseteq \mathfrak{sp}(E)^{(2)} \cong \mathfrak{psl}_4$ . Hence  $\mathfrak{g} \cong \mathfrak{psl}_4$  by dimension reasons.

Set  $V := O(2; \underline{1})$ , a 4-dimensional vector space over  $\mathbb{k}$ , and regard  $V$  as the left regular  $O(2; \underline{1})$ -module. The Lie algebra  $W(2; \underline{1})$  acts on  $V$  by derivations. Let  $\hat{\mathcal{W}}$  denote the semidirect product  $W(2; \underline{1}) \ltimes O(2; \underline{1})$  where  $O(2; \underline{1})$  is regarded as an abelian ideal of  $\hat{\mathcal{W}}$  acted upon by  $W(2; \underline{1})$  in a natural fashion. The above-mentioned actions of  $O(2; \underline{1})$  and  $W(2; \underline{1})$  are compatible in the sense that they give rise to a faithful representation  $\rho: \hat{\mathcal{W}} \rightarrow \mathfrak{gl}(V)$ . It is straightforward to see that the representation  $\rho$  is irreducible.

In characteristic 2, the standard grading of  $L := W(2; \underline{1})$  is surprisingly short, namely,  $L = L_{-1} \oplus L_0 \oplus L_1$ , and we have that  $L_0 \cong \mathfrak{gl}_2$  and  $[L_{-1}, L_1] = L_0$ . Furthermore,  $L_1 \cong L_{-1}^*$  is an irreducible 2-dimensional  $L_0$ -module. From this it is immediate that  $L \cong \mathfrak{sl}_3$  as Lie algebras (this can also be deduced from the fact that  $L$  acts faithfully on the 3-dimensional vector space  $O(2; \underline{1})/\mathbb{k}1$ ). In particular,  $L$  is simple. Let  $\mathfrak{m}$  be the maximal ideal of  $O(2; \underline{1})$ , so that  $O(2; \underline{1}) = \mathbb{k}1 \oplus \mathfrak{m}$ . Obviously,  $\rho(1) = \text{Id}_V$  and  $\rho(\mathfrak{m})^2 = 0$ . Since the trace of  $\text{Id}_V$  is zero and  $W(2; \underline{1}) = W(2; \underline{1})^{(1)}$  by the preceding remark,  $\rho(\hat{\mathcal{W}}) \subset \mathfrak{sl}(V)$ . Note that  $\mathbb{k}1$  coincides with the centre of the Lie algebra  $\hat{\mathcal{W}}$ . We now set  $\mathcal{W} := \hat{\mathcal{W}}/\mathbb{k}1$ , a semidirect product of  $W(2; \underline{1})$  and  $\text{nil}(\mathcal{W}) = O(2; \underline{1})/\mathbb{k}1$ . By the above,  $\mathcal{W}$  is an 11-dimensional Lie subalgebra of  $\mathfrak{psl}(V) \cong \mathfrak{g}$ .

We claim that  $\mathcal{W}$  is a maximal subalgebra of  $\mathfrak{psl}(V)$ . Indeed, suppose the contrary. Then  $\mathfrak{psl}(V)$  has a subalgebra of codimension  $\leq 2$ , say  $\mathfrak{r}$ . Taking the inverse image of  $\mathfrak{r}$  in  $\mathfrak{sl}(V)$  we observe that  $\mathfrak{sl}(V)$  has a proper Lie subalgebra of codimension  $\leq 2$ . Since the set of all  $d$ -dimensional subalgebras of  $\mathfrak{sl}(V)$  is Zariski closed in the Grassmannian  $\text{Gr}(d, \mathfrak{sl}(V))$  and is acted upon by a Borel subgroup  $B$  of  $\text{SL}(V)$ , it follows from the Borel fixed-point theorem that  $\mathfrak{sl}(V)$  contains a Lie subalgebra of codimension  $\leq 2$  normalised by  $\text{Ad } B$ ; we call it  $\mathfrak{p}$ . Since the  $\mathbb{k}$ -span of  $\text{Id}_V$  is the only proper nonzero ideal of  $\mathfrak{sl}(V)$ , the Lie subalgebra  $\mathfrak{p} + \text{Lie}(B)$  of  $\mathfrak{sl}(V)$  is proper and  $(\text{Ad } B)$ -stable. From this it is immediate that  $\text{SL}(V)$  contains a proper parabolic subgroup of codimension  $\leq 2$ . Since  $\dim V = 4$  we reach a contradiction thereby proving the claim.

If  $\mathcal{W}$  is a parabolic subalgebra of  $\mathfrak{g} \cong \mathfrak{psl}(V)$  then its preimage  $\rho(\hat{\mathcal{W}})$  in  $\mathfrak{sl}(V)$  is a proper parabolic subalgebra of  $\mathfrak{sl}(V)$ . But then  $V$  is a reducible  $\rho(\hat{\mathcal{W}})$ -module. As this contradicts the irreducibility

of  $\rho: \hat{\mathcal{W}} \rightarrow \mathfrak{gl}(V)$  we conclude that the subalgebra  $\mathcal{W}$  of  $\mathfrak{g}$  is a counterexample to Morozov's theorem in type  $G_2$ .

#### REFERENCES

- [BGP09] G. Benkart, T. Gregory and A. Premet. The Recognition Theorem for Graded Lie Algebras in Prime Characteristic. *Mem. Amer. Math. Soc.*, 197(920):xii+145 pp., 2009.
- [BIO79] G.M. Benkart, I.M. Isaacs and J.M. Osborn. Albert–Zassenhaus Lie algebras and isomorphisms, *J. Algebra*, 57:310–338, 1979.
- [Bl62] R.E. Block. Trace forms on Lie algebras, *Canad. J. Math.*, 14:553–564, 1962.
- [BW82] R.E. Block and R.L. Wilson. The simple Lie  $p$ -algebras of rank 2, *Ann. of Math.*, 115:93–168, 1982.
- [Bor91] A. Borel. *Linear Algebraic Groups*. Graduate Texts in Mathematics, Vol. 126. Springer-Verlag, New York, second edition, 1991.
- [BT71] A. Borel and J. Tits. Éléments unipotents et sous-groupes paraboliques de groupes réductifs, I, *Invent. Math.* 12:95–104, 1971.
- [Bou68] N. Bourbaki. *Groupes et Algèbres de Lie*, IV, V, VI. Hermann, Paris, 1968.
- [Bou75] N. Bourbaki. *Groupes et Algèbres de Lie*, VII, VIII. Hermann, Paris, 1975.
- [CLNP3] J.F. Carlson, Z. Lin, D.K. Nakano, B.J. Parshall. The restricted nullcone, *Contemp. Math.* 325:51–75, 2003.
- [Car93] R.W. Carter. *Finite Groups of Lie Type: Conjugacy Classes and Complex Characters*. Wiley Classics Library. John Wiley & Sons Ltd., Chichester, 1993. Reprint of the 1985 Wiley-Interscience publication.
- [CE15] A. Castillo-Ramirez and A. Elduque. Some special features of Cayley algebras, and  $G_2$ , in low characteristic. arXiv preprint [arXiv:1503.05529v2](https://arxiv.org/abs/1503.05529v2) [math.RA], 2015.
- [CP13] M.C. Clarke and A. Premet. The Hesselink stratification of nullcones and base change. *Invent. Math.*, 191:631–669, 2013.
- [dGE09] W.A. de Graaf and A. Elashvili. Induced nilpotent orbits of the simple Lie algebras of exceptional type. *Georgian Math. J.*, 16:257–278, 2009.
- [Dyn52] E.B. Dynkin. Semisimple subalgebras of semisimple Lie algebras (Russian). *Matem. Sb. N.S.*, 30(72):349–462, 1952.
- [FSW15] J. Feldvoss, S. Siciliano and Th. Weigel. Restricted Lie algebras with 0-PIM. arXiv preprint [arXiv:1407.1902v1](https://arxiv.org/abs/1407.1902v1) [math.RT], 2014; to appear in *Transform. Groups*.
- [Gar09] S. Garibaldi. Vanishing of trace forms in low characteristics. *Algebra Number Theory*, 3(5):543–566, 2009. With an appendix by Alexander Premet.
- [HS15a] S. Herpel and D. Stewart. On the smoothness of normalisers, the subalgebra structure of modular Lie algebras and the cohomology of small representations. arXiv preprint [arXiv:1402.6280v3](https://arxiv.org/abs/1402.6280v3) [math.GR], 2015.
- [HS15b] S. Herpel and D. Stewart. Maximal subalgebras of Catran type in exceptional Lie algebras. *Selecta Math. (N.S.)*, DOI 10.1007/s00029-015-0199-5; arXiv preprint [arXiv:1409.3862v2](https://arxiv.org/abs/1409.3862v2) [math.RA], 2015.
- [RH98] R.R. Holmes. Simple restricted modules for the restricted Hamiltonian algebra, *J. Algebra*, 199:229–261, 1998.
- [RH01] R.R. Holmes. Simple modules with character height at most one for the restricted Witt algebras, *J. Algebra*, 237:446–469, 2001.
- [Hum67] J.E. Humphreys. Algebraic Groups and Modular Lie Algebras. *Mem. Amer. Math. Soc.*, no. 71, 76 pp., 1967.
- [Ku91] M.I. Kuznetsov. Melikyan algebras as Lie algebras of type  $G_2$ . *Comm. Algebra*, 19:1281–1312, 1991.
- [Law95] R. Lawther. Jordan block sizes of unipotent elements in exceptional algebraic groups. *Comm. Algebra*, 23:4125–4156, 1995.
- [LT07] R. Lawther and D.M. Testerman. Centres of centralizers of unipotent elements in simple algebraic groups, preprint 2007, 301 pp.
- [LT11] R. Lawther and D.M. Testerman. Centres of Centralizers of Unipotent Elements in Simple Algebraic Groups. *Mem. Amer. Math. Soc.*, 210(988):vi+188 pp., 2011.
- [LMT09] P. Levy, G.J. McNinch and D.M. Testerman. Nilpotent subalgebras of semisimple Lie algebras. *C.R. Math. Acad. Sci. Paris*, 347:477–482, 2009.
- [Ma45] A.I. Mal'cev. Commutative subalgebras of semi-simple Lie algebras (Russian). *Bull. Acad. Sci. USSR Sér. Math.* [Izvestia Acad. Nauk SSSR], 9:291–300, 1945.
- [McN03] G.J. McNinch. Sub-principal homomorphisms in positive characteristic. *Math. Z.*, 244:433–455, 2003.
- [Mor56] V.V. Morozov. Proof of the theorem of regularity (Russian). *Uspehi Mat. Nauk*, 11:191–194, 1956.

- [PV10] D.I. Panyushev and E.B. Vinberg. The work of Vladimir Morozov on Lie algebras. *Transform. Groups*, 15:1001–1013, 2010. Special issue dedicated to V.V. Morozov.
- [PeS15] J. Pevtsova and J. Stark. Varieties of elementary subalgebras of maximal dimension for modular Lie algebras. arXiv preprint [arXiv:1503.01043v1 \[math.RT\]](#), 2015.
- [Pre87] A. Premet. On Cartan subalgebras of Lie  $p$ -algebras. *Math. USSR-Izv.*, 29:145–157, 1987.
- [Pre90] A. Premet. Regular Cartan subalgebras and nilpotent elements in restricted Lie algebras. *Math. USSR-Sb.*, 66:555–570, 1990.
- [Pre94] A. Premet. A generalisation of Wilson’s theorem on Cartan subalgebras of simple Lie algebras. *J. Algebra*, 167:641–703, 1994.
- [Pre03] A. Premet. Nilpotent orbits in good characteristic and the Kempf–Rousseau theory. *J. Algebra*, 260:338–366, 2003. Special issue celebrating the 80th birthday of Robert Steinberg.
- [PSk99] A. Premet and S. Skryabin. Representations of restricted Lie algebras and families of associative  $\mathcal{L}$ -algebras. *J. Reine Angew. Math.*, 507:189–218, 1999.
- [PSt15] A. Premet and D. Stewart. Rigid orbits and sheets in reductive Lie algebras over fields of prime characteristic. arXiv preprint [arXiv:1507.05303v2 \[math.RT\]](#), 2015.
- [PS99] A. Premet and H. Strade. Simple Lie algebras of small characteristic: II. Exceptional roots. *J. Algebra*, 216:190–301, 1999.
- [PS01] A. Premet and H. Strade. Simple Lie algebras of small characteristic: III. The toral rank 2 case. *J. Algebra*, 242:236–337, 2001.
- [PS08] A. Premet and H. Strade. Simple Lie algebras of small characteristic: VI. Completion of the classification. *J. Algebra*, 320:3559–3604, 2008.
- [PSu83] A. Premet and I.D. Suprunenko. Quadratic modules for Chevalley groups over fields of odd characteristic. *Math. Nachr.*, 110:65–96, 1983.
- [Sel67] G. Seligman. *Modular Lie Algebras*. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 40, Springer-Verlag, New York, 1967.
- [Ser06] J.-P. Serre. Coordonnées de Kac. *Oberwolfach Reports*, 3:1787–1790, 2006.
- [Skr91] S.M. Skryabin. Modular Lie algebras of Cartan type over algebraically non-closed fields, I. *Comm. Algebra*, 19:1629–1741, 1991.
- [Skr98] S.M. Skryabin. Toral rank one simple Lie algebras of low characteristic. *J. Algebra*, 200:650–700, 1998.
- [Spr66] T.A. Springer. Some arithmetic results on semi-simple Lie algebras. *Inst. Hautes Études Sci. Publ. Math.*, 30:115–141, 1966.
- [St61] R. Steinberg. Automorphisms of classical Lie algebras. *Pacific J. Math.*, 11:1119–1129, 1961.
- [Str04] H. Strade. *Simple Lie Algebras over Fields of Positive Characteristic. I. Structure Theory*. de Gruyter Expositions in Mathematics, Vol. 38, Walter de Gruyter & Co., Berlin, 2004.
- [Str09] H. Strade. *Simple Lie Algebras over Fields of Positive Characteristic. II. Classifying the Absolute Toral Rank Two Case*. de Gruyter Expositions in Mathematics, Vol. 42, Walter de Gruyter & Co., Berlin, 2009.
- [VAG05] University of Georgia VIGRE Algebra Group. Varieties of nilpotent elements for simple Lie algebras: II. Bad primes. *J. Algebra*, 292:65–99, 2005. The University of Georgia VIGRE Algebra Group: D.J. Benson, P. Bergonio, B.D. Boe, L. Chastkofsky, B. Cooper, G.M. Guy, J. Hower, M. Hunziker, J.J. Hyun, J. Kujawa, G. Matthews, N. Mazza, D.K. Nakano, K.J. Platt and C. Wright.
- [Weis66] B.Ju. Weisfeiler. A class of unipotent subgroups of semisimple algebraic groups (Russian). *Uspehi Mat. Nauk*, 21:222–223, 1966. [arXiv:math/0005149v1 \[math.AG\]](#) (English translation).
- [Weis78] B.Ju. Weisfeiler. On the structure of the minimal ideal of some graded Lie algebras of characteristic  $p > 0$ . *J. Algebra*, 53:344–361, 1978.

SCHOOL OF MATHEMATICS, THE UNIVERSITY OF MANCHESTER, OXFORD ROAD, M13 9PL, UK

*E-mail address:* [alexander.premet@manchester.ac.uk](mailto:alexander.premet@manchester.ac.uk)